Lecture Notes on Differential Games

Alberto Bressan
Department of Mathematics, Penn State University,
University Park, Pa. 16802, USA.
bressan@math.psu.edu

November 3, 2010

Abstract

These notes provide a brief introduction to some aspects of game theory.

1 Introduction

A basic problem in optimization theory is to find the maximum value of a function:

$$\max_{x\in X}\ \phi(x)\,. \tag{1.1}$$

Typically, ϕ is a continuous function and the maximum is sought over a closed, possibly unbounded domain $X \subseteq I\!R^m$. An extensive mathematical theory is currently available on the existence of the maximum, on necessary and sufficient conditions for optimality, and on computational methods. Interpreting ϕ as a payoff function, one can regard (1.1) as a decision problem. Among all possible choices $x \in X$, we seek the one that provides the maximum possible payoff.

As in (1.1), **optimization theory** deals with the case where there is only one individual, making a decision and achieving a payoff. **Game theory**, on the other hand, is concerned with the more complex situation where two or more individuals, or "players" are present. Each player can choose among a set of available options. His payoff, however, depends also on the choices made by all the other players.

For simplicity, consider the case of two players. Player 1 can choose a strategy $x_1 \in X_1$, while Player 2 can choose $x_2 \in X_2$. For $i = 1, 2$, the goal of Player i is

$$\text{maximize:}\quad \phi_i(x_1, x_2)\,. \tag{1.2}$$

In contrast with (1.1), it is clear that the problem (1.2) does not admit an "optimal" solution. Indeed, in general it will not be possible to find a couple $(\bar x_1, \bar x_2) \in X_1 \times X_2$ which at the same time maximizes the payoff of the first player and of the second player, so that

$$\phi_1(\bar x_1, \bar x_2) \;=\; \max_{x_1,x_2}\ \phi_1(x_1,x_2)\,, \qquad\qquad \phi_2(\bar x_1, \bar x_2) \;=\; \max_{x_1,x_2}\ \phi_2(x_1,x_2)\,.$$

For this reason, various alternative concepts of solutions have been proposed in the literature. These can be relevant in different situations, depending on the information available to the players and their ability to cooperate.

For example, if the players have no means to talk to each other and do not cooperate, then an appropriate concept of solution is the Nash equilibrium, defined as a fixed point of the best reply map. In other words, (x_1^*, x_2^*) is a Nash equilibrium if

(i) the value $x_1^* \in X_1$ is the best choice for the first player, in reply to the strategy x_2^* adopted by the second player. Namely

$$\phi_1(x_1^*, x_2^*) \;=\; \max_{x_1 \in X_1} \; \phi_1(x_1, x_2^*),$$

(ii) the value $x_2^* \in X_2$ is the best choice for the second player, in reply to the strategy x_1^* adopted by the first player. Namely

$$\phi_2(x_1^*, x_2^*) \;=\; \max_{x_2 \in X_2} \; \phi_2(x_1^*, x_2).$$

On the other hand, if the players can cooperate and agree on a joint course of action, their best strategy $(x_1^*, x_2^*) \in X_1 \times X_2$ will be one which maximizes the sum:

$$\phi_1(x_1^*, x_2^*) + \phi_2(x_1^*, x_2^*) \;=\; \max_{x_1, x_2} \; \Big[\phi_1(x_1, x_2) + \phi_2(x_1, x_2)\Big] .$$

In general, in order to be acceptable to both players, this strategy will also require a side payment to compensate the player with the smaller payoff.

The situation modeled by (1.2) represents a *static game*, sometimes also called a "one-shot" game. Each player makes one choice $x_i \in X_i$, and this completely determines the payoffs. In other relevant situations, the game takes place not instantaneously but over a whole interval of time. This leads to the study of *dynamic games*, also called "evolutionary games". In this case, the strategy adopted by each player is described by a function of time $t \mapsto u_i(t)$. Here the time variable t can take a discrete set of values, or range over a whole interval $[0, T]$.

We recall that, in the standard model of **control theory**, the state of a system is described by a variable $x \in I\!R^n$. This state evolves in time, according to an ODE

$$\dot x(t) \;=\; f(t, x(t),\, u(t)) \qquad\qquad t \in [0, T] \,. \tag{1.3}$$

Here $t \mapsto u(t) \in U$ is the *control function*, ranging within a set U of admissible control values.

Given an initial condition

$$x(0) = x_0 \,, \tag{1.4}$$

a basic problem in optimal control is to find a control function $u(\cdot)$ which maximizes the payoff

$$J(u) \;=\; \psi(x(T)) - \int_0^T L(t, x(t),\, u(t))\, dt \,. \tag{1.5}$$

Here ψ is a *terminal payoff*, while L accounts for a *running cost*.

Differential games provide a natural extension of this model to the case where two or more individuals are present, and each one of them seeks to maximize his own payoff. In the case of two players, one thus considers a system whose state $x \in I\!R^n$ evolves according to the ODE

$$\dot{x}(t) \;=\; f(t, x(t),\, u_1(t),\, u_2(t)) \qquad\qquad t \in [0, T]\,. \tag{1.6}$$

Here $t \mapsto u_i(t) \in U_i$, $i = 1, 2$, are the *control functions* implemented by the two players.

Given the initial condition (1.4), the goal of the i-th player is

$$\text{maximize:} \qquad J_i \;=\; \psi_i(x(T)) - \int_0^T L_i(t, x(t),\, u_1(t),\, u_2(t))\, dt\,. \tag{1.7}$$

As in the case of one-shot games, various concepts of solution can be considered. In addition, one can further distinguish between *open-loop* strategies $u_i = u_i(t)$, depending only on the time variable, and *feedback strategies* $u_i = u_i(t, x)$, depending also on the current state of the system. In a situation where each player has knowledge only of the initial state of the system, it is natural to consider open-loop strategies. On the other hand, if the players can observe the current state of the system, it is more appropriate to consider feedback strategies.

In the literature, a first, well known example of a non-cooperative game in economics appeared in [12]. Within this monograph, Cournot studied a duopoly, where two firms selling the same product seek to adjust their production levels in order to maximize profits. His solution can be interpreted as the fixed point of a best reply map.

The classic book [27] by von Neumann and Morgenstern is widely regarded as the starting point of the mathematical theory of games. While this book focuses on two-players, zero-sum games, the later paper of Nash [21] provided a concept of solution for general non-cooperative games for N players.

The theory of differential games was first developed by Isaacs [18], followed by other authors; see [16, 20]. A comprehensive presentation of dynamic games, with applications to economic models, can be found in [5, 13].

Aim of the present notes is to provide a concise introduction to the mathematical theory of games for two players. The first chapter deals with static games, while the remaining chapters deal with dynamic games.

For static games, the existence of Nash equilibrium solutions is proved by an application of the Kakutani fixed point theorem for multivalued maps. Using the approximate selection theorem of Cellina, this can be derived as an easy consequence of the classical Brouwer fixed point theorem. Specializing to zero-sum games, some basic results by von Neumann can then be deduced as corollaries.

The analysis of differential games relies heavily on concepts and techniques of optimal control theory. Equilibrium strategies in open-loop form can be found by solving a two-point boundary value problem for an ODE derived from the Pontryagin maximum principle. On the other hand, equilibrium strategies in feedback form are best studied by looking at a system of Hamilton-Jacobi-Bellman PDEs for the value functions of the various players, derived from the principle of dynamic programming.

A review of background material on multifunctions, fixed point theorems, and control theory, is provided in the Appendix to these lecture notes.

2 Static Games

In its 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: $\Phi^A : A \times B \mapsto I\!R$ and $\Phi^B : A \times B \mapsto I\!R$.

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**. Clearly, a zero-sum game is determined by one single payoff function $\Phi = \Phi^A = -\Phi^B$.

Throughout the following, our basic assumption will be

(A1) The sets A and B are compact metric spaces. The payoff functions Φ^A, Φ^B are continuous functions from $A \times B$ into $I\!R$.

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

$$A \doteq \{a_1, a_2, \ \ldots\ , a_m\}, \qquad\qquad B \doteq \{b_1, b_2, \ \ldots\ , b_n\}. \tag{2.1}$$

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

$$\Phi^A_{ij} \doteq \Phi^A(a_i, b_j), \qquad\qquad \Phi^B_{ij} \doteq \Phi^B(a_i, b_j). \tag{2.2}$$

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 two numbers: Φ^A_{ij}, Φ^B_{ij}, see figures 3, 4. 5.

2.1 Solution concepts

In general, one cannot speak of an "optimal solution" of the game. Indeed, an outcome that is optimal for one player can be very bad for the other one. We review here 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. *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 other words, it is not possible to strictly increase the payoff of one player without strictly decreasing the payoff of the other.

In general, a game can admit several Pareto optima (see Fig. 6). In order to construct a pair of strategies which is Pareto optimal, one can proceed as follows. Choose any number $\lambda \in [0,1]$ and consider the optimization problem

$$\max_{(a,b)\in A\times B} \quad \lambda\,\Phi^A(a,b) + (1-\lambda)\,\Phi^B(a,b)\,. \tag{2.3}$$

By the compactness and continuity assumptions (A1), an optimal solution does 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:

$$R^A(b) \doteq \Big\{a \in A\,;\;\; \Phi^A(a,b) \;=\; \max_{\omega\in A}\Phi^A(\omega,b)\Big\}\,. \tag{2.4}$$

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

$$R^B(a) \doteq \Big\{b \in B\,;\;\; \Phi^B(a,b) \;=\; \max_{\omega\in B}\Phi^B(a,\omega)\Big\}\,. \tag{2.5}$$

By the assumption (A1), 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 strategy in advance, and then player B (the follower) makes his choice accordingly.

In this case, the game can be reduced to a pair of 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 payoff function $b \mapsto \Phi^B(a,b)$. 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 he will choose the one which is most favorable to Player A.

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 moreover*

$$\Phi^A(a,b) \ \leq \ \Phi^A(a_S, b_S) \qquad \text{for every pair } (a,b) \text{ with } b \in R^B(a).$$

Under the assumption (A1), it is easy to check that a Stackelberg equilibrium always exists. Indeed, consider the domain

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

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 $(a_S, b_S) \in R$. This yields a Stackelberg equilibrium.

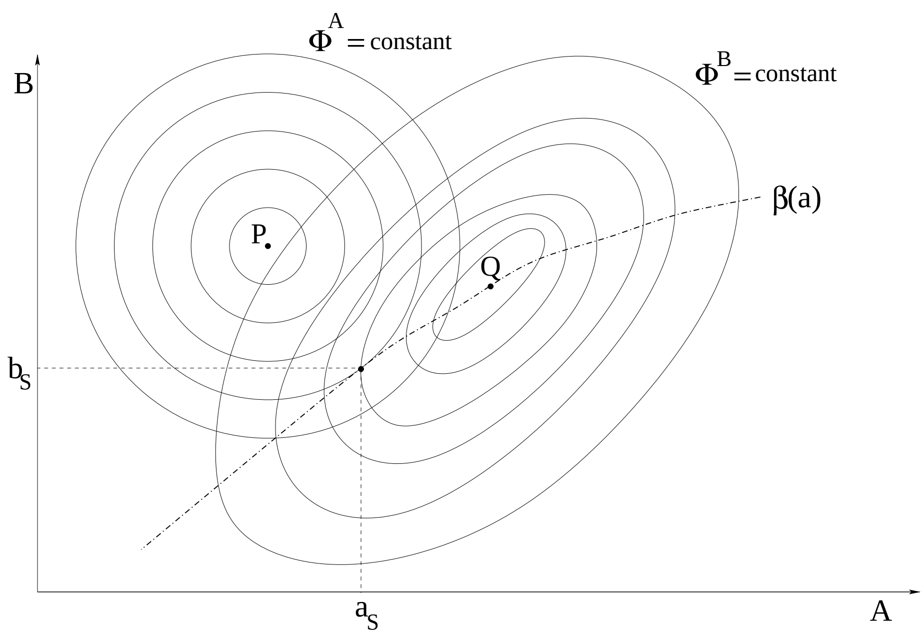

Figure 1: The figure shows the level curves of the two payoff functions. Here player A chooses the horizontal coordinate, player B the vertical coordinate. The payoff function Φ^A attains its global maximum at P, while Φ^B attains its maximum at Q. If the first player chooses a strategy $a \in A$, then $\beta(a) \in B$ is the best reply for the second player. The pair of strategies $(a_S\,,\ b_S)$ is a Stackelberg equilibrium. Notice that at this point the curve $b = \beta(a)$ is tangent to a level curve of Φ^A.

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.

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*

$$\Phi^A(a, b^*) \ \leq \ \Phi^A(a^*, b^*)\,, \qquad\qquad \Phi^B(a^*, b) \ \leq \ \Phi^B(a^*, b^*)\,. \tag{2.6}$$

In other words, no player can increase his payoff by single-mindedly changing his strategy, as long as the other player sticks to the 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:

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

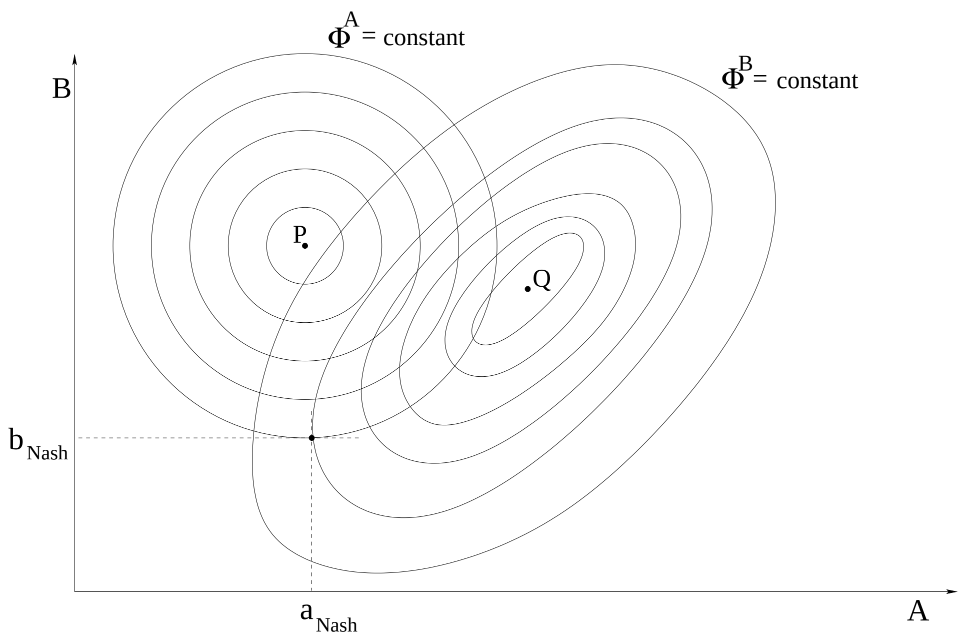

Figure 2: Here player A chooses the horizontal coordinate, player B the vertical cooordinate. The payoff function Φ^A attains its global maximum at P, while Φ^B attains its global maximum at Q. The pair of strategies $(a_{\rm Nash}\,,\ b_{\rm Nash})$ is a Nash equilibrium. Notice that at this point the level curve of Φ^A has horizontal tangent while the level curve of Φ^B has vertical tangent.

The following examples show that:

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

(ii) the 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 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 Figure 3. By direct inspection, one checks that it does not admit any Nash equilibrium solution.

Player B

Player A	H	T
H	1, −1	−1, 1
T	−1, 1	1, −1

	H	T
H	1	−1
T	−1	1

Figure 3: 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 2. Consider the game whose bi-matrix of payoffs is given in Figure 4. The pair of strategies (a_1, b_3) is a Nash equilibrium, as well as a Pareto optimum. On the other hand, the pair of strategies (a_2, b_1) is a Nash equilibrium but not a Pareto optimum. Indeed, (a_1, b_3) is the unique Pareto optimum.

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

Figure 4: A bi-matrix of payoffs, with two Nash equlibrium points but only one Pareto optimum.

Example 3 (prisoners' dilemma). Consider the game with payoffs described by the bi-matrix in Figure 5. 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 he confesses, the police rewards him by reducing his sentence. None of the prisoners, while interrogated, knows about the behavior of the other.

Player B

Player A	C	N
C	−6, −6	0, −8
N	−8, 0	−1, −1

Figure 5: The bi-matrix of payoffs for the "prisoners' dilemma".

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, my 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 my two options result in either 0 or 1 years in jail. Again, confessing is the best choice. Since the 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. 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.

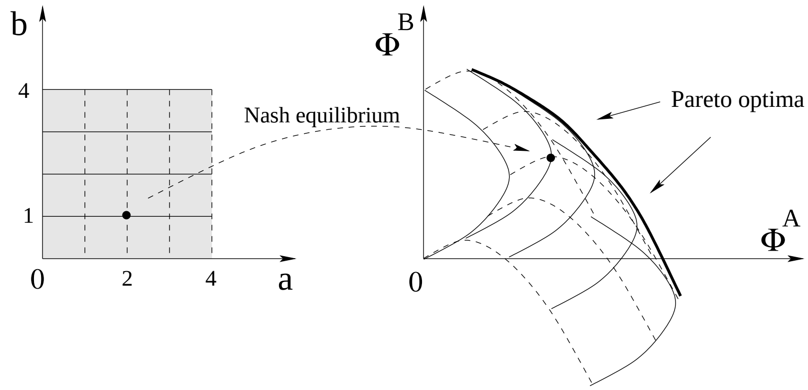

Figure 6: The payoff map $(a, b) \mapsto (\Phi^A, \Phi^B)$ for the game (2.7). The Nash equilibrium is $(2, 1)$, which does not yield a Pareto optimum.

Example 4. Let $A = B = [0, 4]$ and consider the payoff functions (see figure 6)

$$\Phi^A(a,b) \;=\; 2a + 2b - \frac{a^2}{2}\,, \qquad\qquad \Phi^B(a,b) \;=\; a + b - \frac{b^2}{2}\,. \tag{2.7}$$

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

$$a^* \;=\; \underset{a\in A}{\operatorname{argmax}} \left\{2a + 2b^* - \frac{a^2}{2}\right\} \;=\; 2\,,$$

$$b^* \;=\; \underset{b\in B}{\operatorname{argmax}} \left\{a^* + b - \frac{b^2}{2}\right\} \;=\; 1\,.$$

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

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

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

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

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_λ, b_λ), with

$$a_\lambda \;=\; \arg\max_{a\in[0,4]} \left\{(\lambda+1)a - \frac{\lambda a^2}{2}\right\} \;=\; \min\left\{1 + \frac{1}{\lambda}\,,\;\; 4\right\},$$

$$b_\lambda \;=\; \arg\max_{b\in[0,4]} \left\{(\lambda+1)b - \frac{(1-\lambda)b^2}{2}\right\} \;=\; \min\left\{\frac{1+\lambda}{1-\lambda}\,,\;\; 4\right\}.$$

2.2 Existence of Nash equilibria

We now state a basic existence theorem for a Nash equilibrium, valid under suitable continuity and convexity assumptions. The proof is a straightforward application of Kakutani's fixed point theorem.

Theorem 1 (existence of Nash equilibria). *Assume that the sets of strategies A, B are compact, convex subsets of $\mathbb{R}^n$. Let the payoff functions Φ^A, Φ^B be continuous and assume that*

$$a \mapsto \Phi^A(a,b) \text{ is a concave function of } a, \text{ for each fixed } b \in B,$$
$$b \mapsto \Phi^B(a,b) \text{ is a concave function of } b, \text{ for each fixed } a \in A.$$

Then the non-cooperative game admits a Nash equilibrium.

Proof. Consider the best reply maps R^A, R^B, defined at (2.4)-(2.5).

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

$$a \mapsto m(a) \doteq \max_{b\in B} \Phi^B(a,b)$$

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

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

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. □

2.3 Randomized strategies

If the convexity 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.

As shown by Example 1, 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. *A* **randomized strategy** *for player A is a probability distribution μ on the his set of strategies A. Similarly, a randomized strategy for player B is a probability distribution ν on the set B.*

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

$$\widetilde{\Phi}^A(\mu,\nu) \doteq \int_{A\times B} \Phi^A(a,b)\ d\mu\otimes d\nu\,. \qquad\qquad \widetilde{\Phi}^B(\mu,\nu) \doteq \int_{A\times B} \Phi^B(a,b)\ d\mu\otimes d\nu\,. \tag{2.8}$$

Remark 1. The above quantities $\widetilde{\Phi}^A(\mu,\nu)$ and $\widetilde{\Phi}^B(\mu,\nu)$ are the **expected values of the payoffs**, if the two players choose random strategies, independent of each other, according to the probability distributions μ, ν, respectively.

In the following, by $\mathcal{P}(A), \mathcal{P}(B)$ 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 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 $x = (x_1, \ldots, x_m) \in \Delta_m$, where

$$\Delta_m \doteq \left\{ x = (x_1,\ldots,x_m)\,; \qquad x_i \in [0,1]\,, \quad \sum_{i=1}^m x_i = 1 \right\}. \tag{2.9}$$

Here x_i is the probability that player A chooses the strategy a_i.

Given the bi-matrix game described at (2.1)-(2.2), the corresponding randomized game can be represented as follows. The two players choose from the sets of strategies

$$\widetilde{A} \doteq \Delta_m\,, \qquad\qquad \widetilde{B} \doteq \Delta_n\,. \tag{2.10}$$

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

$$\widetilde{\Phi}^A(x,y) \doteq \sum_{ij} \Phi^A_{ij}\ x_i y_j\,, \qquad\qquad \widetilde{\Phi}^B(x,y) \doteq \sum_{ij} \Phi^B_{ij}\ x_i y_j\,. \tag{2.11}$$

The concept of Nash equilibrium admits a natural extension to the class of randomized strategies. A fundamental result proved by J. Nash is that every game has an equilibrium solution, within the family of randomized strategies.

Theorem 2 (existence of Nash equilibria for randomized strategies). *Let the assumptions (A1) hold. Then there exist probability measures $\mu^* \in \mathcal{P}(A)$ and $\nu^* \in \mathcal{P}(B)$ such that*

$$\widetilde{\Phi}^A(\mu, \nu^*) \leq \widetilde{\Phi}^A(\mu^*, \nu^*) \qquad \textit{for all} \quad \mu \in \mathcal{P}(A)\,, \tag{2.12}$$

$$\widetilde{\Phi}^B(\mu^*, \nu) \leq \widetilde{\Phi}^B(\mu^*, \nu^*) \qquad \textit{for all} \quad \nu \in \mathcal{P}(B)\,. \tag{2.13}$$

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 (2.1)-(2.2). We check that all assumptions of Theorem 1 are satisfied.

The sets of randomized strategies, defined at (2.9)-(2.10), are compact convex simplexes. The payoff functions $\Phi^A, \Phi^B : \Delta_m \times \Delta_n \mapsto I\!R$, defined at (2.11), are bilinear, hence continuous.

For each given strategy $y \in \Delta_n$ chosen by the second player, the payoff function for the first player

$$x \mapsto \Phi^A(x, y) = \sum_{i,j} \Phi^A_{ij}\, x_i y_j$$

is linear, hence concave. Similarly, for each $x \in \Delta_m$, the payoff function for the second player

$$y \mapsto \Phi^B(x, y) = \sum_{i,j} \Phi^B_{ij}\, x_i y_j$$

is linear, hence concave.

We can thus apply Theorem 1 and obtain the existence of a Nash equilibrium solution $(x^*, y^*) \in \Delta_m \times \Delta_n$.

2. In the remainder of the proof, using an approximation argument we extend the result to the general case where A, B are compact metric spaces. Let $\{a_1, a_2, \ldots\}$ be a sequence of points dense in A, and let $\{b_1, b_2, \ldots\}$ 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 $A_n \doteq \{a_1, \ldots, a_n\}$ and $B_n \doteq \{b_1, \ldots, b_n\}$. By the previous step, this game has a Nash equilibrium solution, given by a pair of randomized strategies (μ_n, ν_n). Here μ_n and ν_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

$$\mu_n \rightharpoonup \mu^*, \qquad \nu_n \rightharpoonup \nu^* \qquad \text{as} \quad n \to \infty, \tag{2.14}$$

for some probability measures $\mu^* \in \mathcal{P}(A)$ and $\nu^* \in \mathcal{P}(B)$.

3. We claim that the pair (μ^*, ν^*) in (2.14) provides a Nash equilibrium solution, i.e. (2.12)-(2.13) hold. This will be proved by showing that

$$\int_{A\times B} \Phi^A(a,b)\ d\mu^* \otimes d\nu^* \ = \ \max_{\mu\in\mathcal{P}(A)} \int_{A\times B} \Phi^A(a,b)\ d\mu \otimes d\nu^*, \tag{2.15}$$

together with the analogous property for Φ^B.

Let $\varepsilon > 0$ be given. By the assumption (A1), there exists $\delta > 0$ such that

$$d(a,a') \le \delta \quad \text{and} \quad d(b,b') \le \delta \qquad \text{imply} \qquad \Big|\Phi^A(a,b) - \Phi^A(a',b')\Big| \ < \ \varepsilon. \tag{2.16}$$

Since the sequences $\{a_k\,;\ \ k \ge 1\}$ and $\{b_k\,;\ \ k \ge 1\}$ 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 $\{\psi_1,\ldots,\psi_N\}$ be a continuous partition of unity on B, subordinated to the covering $\{B(b_j,\delta)\,;\ \ j = 1,\ldots,N\}$.

Any probability measure $\mu \in \mathcal{P}(A)$ 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

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

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

$$\hat\nu(\{b_j\}) \doteq \int \psi_j\, d\nu \qquad\qquad j = 1,\ldots,N\,.$$

For every pair of probability measures (μ,ν), by (2.16) the above construction yields

$$\begin{aligned} &\left|\int_{A\times B} \Phi^A(a,b)\ d\mu\otimes d\nu\ - \int_{A\times B} \Phi^A(a,b)\ d\hat\mu\otimes d\hat\nu\right| \\ &\le\ \int_{A\times B} \sum_{i,j} \varphi_i(a_i)\psi_j(b_j)\,\Big|\Phi^A(a,b) - \Phi^A(a_i,b_j)\Big|\ d\mu\otimes d\nu \\ &\le\ \int_{A\times B} \varepsilon\ d\mu\otimes d\nu\ =\ \varepsilon\,. \end{aligned} \tag{2.17}$$

4. For all $i,j = 1,\ldots,N$, as $n\to\infty$ the weak convergence (2.14) yields

$$\hat\mu_n(\{a_i\})\ =\ \int \varphi_i\, d\mu_n\ \to\ \int \varphi_i\, d\mu^*\ =\ \hat\mu^*(\{a_i\})\,. \tag{2.18}$$

Similarly, $\hat\nu_n(\{b_j\}) \to \hat\nu^*(\{b_j\})$.

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

$$\widetilde{\Phi}^A(\hat{\mu}, \nu_n) \;\leq\; \widetilde{\Phi}^A(\mu_n, \nu_n). \tag{2.19}$$

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, ν_n) provides a Nash equilibrium to the game restricted to $A_n \times B_n$. Using (2.17), (2.18), and (2.19), for every $\mu \in \mathcal{P}(A)$ we obtain

$$\widetilde{\Phi}^A(\mu, \nu^*) - \varepsilon \;\leq\; \widetilde{\Phi}^A(\hat{\mu}, \hat{\nu}^*) \;=\; \lim_{n\to\infty} \widetilde{\Phi}^A(\hat{\mu}, \hat{\nu}_n)$$

$$\leq\; \limsup_{n\to\infty}\; \widetilde{\Phi}^A(\hat{\mu}, \nu_n) + \varepsilon \;\leq\; \lim_{n\to\infty} \widetilde{\Phi}^A(\mu_n, \nu_n) + \varepsilon \;=\; \widetilde{\Phi}^A(\mu^*, \nu^*) + \varepsilon\,.$$

Since $\mu \in \mathcal{P}(A)$ and $\varepsilon > 0$ were arbitrary, this proves (2.12). The proof of (2.13) is entirely similar. □

2.4 Zero-sum games

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

$$\Phi : A \times B \mapsto I\!R\,. \tag{2.20}$$

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

(A1′) The domains A, B are compact metric spaces and the function $\Phi : A \times B \mapsto I\!R$ is continuous.

In particular, this implies that the maps

$$b \mapsto \max_{a\in A}\; \Phi(a, b), \qquad\qquad a \mapsto \min_{b\in B}\; \Phi(a, b) \tag{2.21}$$

are both continuous.

In a symmetric situation, each of the two players will have to make his 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 choice, depending on b.

This is clearly a situation where player A has the advantage of knowing his 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

$$V^+ \doteq \min_{b\in B} \Phi(\alpha(b),\, b) = \min_{b\in B} \max_{a\in A} \Phi(a,b)\,. \tag{2.22}$$

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

$$V^- \doteq \max_{a\in A} \Phi(a,\, \beta(a)) = \max_{a\in A} \min_{b\in B} \Phi(a,b)\,. \tag{2.23}$$

Lemma 1. *In the above setting, one has*

$$V^- \doteq \max_{a\in A} \min_{b\in B} \Phi(a,b) \le \min_{b\in B} \max_{a\in A} \Phi(a,b) \doteq V^+\,. \tag{2.24}$$

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 $\varepsilon > 0$ there exists $a_\varepsilon \in A$ such that

$$\Phi(a_\varepsilon,\, \beta(a_\varepsilon)) > V^- - \varepsilon\,. \tag{2.25}$$

In turn, this implies

$$V^+ = \min_{b\in B} \max_{a\in A} \Phi(a,b) \ge \min_{b\in B} \Phi(a_\varepsilon, b) = \Phi(a_\varepsilon,\, \beta(a_\varepsilon)) > V^- - \varepsilon\,.$$

Since $\varepsilon > 0$ 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 $V \doteq V^- = V^+$ is the **value of the game**.

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

$$\min_{b\in B} \Phi(a^*, b) = \Phi(a^*, b^*) = \max_{a\in A} \Phi(a, b^*)\,, \tag{2.26}$$

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 (2.26), 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 (value and saddle point). *Under the assumptions (A′), the zero-sum game (2.20) has a value V if and only if a saddle point (a^*, b^*) exists. In the positive case, one has*

$$V \;=\; V^- \;=\; V^+ \;=\; \Phi(a^*, b^*)\,. \tag{2.27}$$

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

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

By (2.24) this implies $V \doteq V^- = V^+$, showing that the game has a value.

2. Next, assume $V \doteq V^- = V^+$. Let $a \mapsto \beta(a)$ be the best reply map for player B. For each $\varepsilon > 0$ choose $a_\varepsilon \in A$ such that (2.25) holds. Since the sets A and B are compact, we can choose a subsequence $\varepsilon_n \to 0$ such that the corresponding strategies converge, say

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

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

$$\Phi(a^*, b^*) \;=\; \lim_{n\to\infty} \Phi(a_{\varepsilon_n}\,,\;\beta(a_{\varepsilon_n})).$$

From

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

letting $\varepsilon \to 0$ we conclude

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

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

Remark 3. As noted in Example 2, 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

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

By applying Theorem 1 to the particular case of a zero-sum game we obtain

Corollary 1 (Existence of a saddle point). *Consider a zero-sum game, satisfying the conditions (A1′). Assume that the sets A, B are convex, and moreover*

$a \mapsto \Phi(a,b)$ is a concave function of a, for each fixed $b \in B$,

$b \mapsto \Phi(a,b)$ is a convex function of b, for each fixed $a \in A$.

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

More generally, as stated in Theorem 2, 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 2 (Existence of a saddle point in the class of randomized strategies). *Under the assumptions (A1′), 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^*, ν^*) of probability measures on A and B respectively, such that

$$\int_{A\times B} \Phi(a,b)\, d\mu \otimes d\nu^* \;\leq\; \int_{A\times B} \Phi(a,b)\; d\mu^* \otimes d\nu^* \;\leq\; \int_{A\times B} \Phi(a,b)\; d\mu^* \otimes d\nu\,,$$

for every other probability measures $\mu \in \mathcal{P}(A)$ and $\nu \in \mathcal{P}(B)$.

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 $A = \{a_1, \ldots, a_m\}$ and $B = \{b_1, \ldots, b_n\}$. The sets of randomized strategies can now be identified with the simplexes Δ_m, Δ_n defined at (2.9). Let $\Phi_{ij} \doteq \Phi(a_i, b_j)$. According to Corollary 2, there exist $x^* \in \Delta_m$ and $y^* \in \Delta_n$ such that

$$\max_{x\in\Delta_m} \left(\sum_{i,j} \Phi_{ij} x_i y_j^*\right) \;\leq\; \sum_{i,j} \Phi_{ij}\, x_i^* y_j^* \;\leq\; \min_{y\in\Delta_n} \left(\sum_{i,j} \Phi_{ij}\, x_i^* y_j\right).$$

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

$$\max_{x\in\Delta_m} \left(\sum_{i,j} \Phi_{ij}\, x_i y_j\right) = \max_{i\in\{1,\ldots,m\}} \left(\sum_{j} \Phi_{ij}\, y_j\right),$$

$$\min_{y\in\Delta_n} \left(\sum_{i,j} \Phi_{ij}\, x_i y_j\right) = \min_{j\in\{1,\ldots,m\}} \left(\sum_{i} \Phi_{ij}\, x_i\right).$$

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

$$x \;\mapsto\; \Phi^{min}(x) \;\doteq\; \min_j \left(\sum_i \Phi_{ij}\, x_i\right) \tag{2.28}$$

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

$$y \;\mapsto\; \Phi^{max}(y) \;\doteq\; \max_i \left(\sum_j \Phi_{ij}\, y_j\right) \tag{2.29}$$

attains its global minimum.

Example 5 (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 Figure 7. 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 A \ Player B	**R**	**P**	**S**
R	0	−1	1
P	1	0	−1
S	−1	1	0

Figure 7: The matrix describing the "rock-paper-scissors" game. Its entries represent the payments from player B to player A.

Example 6. 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.

Player A \ Player B	b_1	b_2
a_1	0	3
a_2	1	0

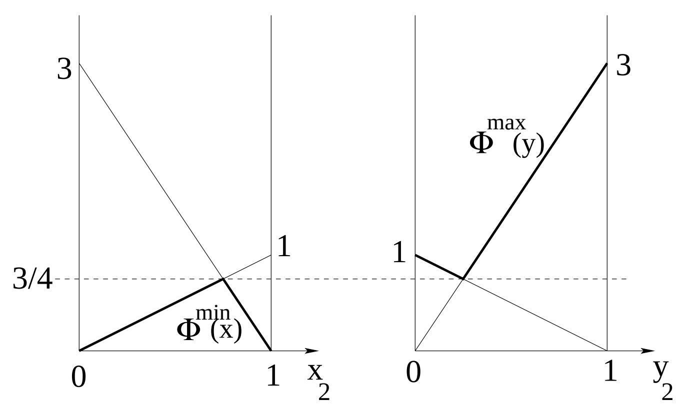

Figure 8: Left: the payoff matrix for the zero-sum game described in Example 6. Center and right: the thick lines represent the graphs of the functions Φ^{min} and Φ^{max}.

This situation can be modeled as a zero-sum game, where the payoff matrix is given in Figure 8, left. Still in Fig. 8, center and right, are the graphs of the functions in (2.28)-(2.29), namely

$$\Phi^{min}(x) = \min_{j=1,2} \left(\Phi_{1j}\, x_1 + \Phi_{2j}\, x_2\right) = \min \left\{3x_1\,,\ x_2\right\}, \qquad (x_1 = 1 - x_2),$$

$$\Phi^{max}(y) = \max_{i=1,2} \left(\Phi_{i1}\, y_1 + \Phi_{i2}\, y_2\right) = \max \left\{y_1\,,\ 3y_2\right\}, \qquad (y_1 = 1 - y_2).$$

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.

2.5 The co-co solution

Consider again a general non-zero sum game, described by the payoff functions

$$\Phi^A : A \times B \mapsto I\!R, \qquad \Phi^B : A \times B \mapsto I\!R, \tag{2.30}$$

under the assumptions (A1). If the two players can talk to each other and cooperate, they can adopt a pair of strategies $(a^\sharp, b^\sharp)$ which maximizes their combined payoffs:

$$V^\sharp \doteq \Phi^A(a^\sharp, b^\sharp) + \Phi^B(a^\sharp, b^\sharp) = \max_{(a,b)\in A\times B} \Big\{\Phi^A(a,b) + \Phi^B(a,b)\Big\}. \tag{2.31}$$

This choice, however, may favor one player much more that the other. For example, one may have $\Phi^B(a^\sharp, b^\sharp) \ll \Phi^A(a^\sharp, b^\sharp)$, an outcome which may not be agreeable to Player B. In this case Player A needs to provide some incentive, in the form of a side payment, inducing Player B to adopt the strategy $b^\sharp$.

In general, splitting the total payoff $V^\sharp$ in two equal parts will not be acceptable, because it does not reflect the relative strength of the players and their personal contributions to the common achievement. A more realistic procedure to split the total payoff among the two players, recently proposed in [19], goes as follows.

Given the two payoff functions Φ^A, Φ^B, define

$$\Phi^\sharp(a,b) \doteq \frac{\Phi^A(a,b) + \Phi^B(a,b)}{2}, \qquad \Phi^\flat(a,b) \doteq \frac{\Phi^A(a,b) - \Phi^B(a,b)}{2}. \tag{2.32}$$

Observing that

$$\Phi^A = \Phi^\sharp + \Phi^\flat, \qquad \Phi^B = \Phi^\sharp - \Phi^\flat,$$

we can split the original game as the sum of a purely cooperative game, where both players have exactly the same payoff $\Phi^\sharp$, and a purely competitive (i.e., zero-sum) game, where the players have opposite payoffs: $\Phi^\flat$ and $-\Phi^\flat$.

Let $V^\sharp$ be as in (2.31) and let $V^\flat$ be the value of the zero-sum game. This value is always well defined, possibly in terms of randomized strategies. The **cooperative-competitive value** (or **co-co value**, in short) of the original game (2.30) is then defined as the pair of payoffs

$$\left(\frac{V^\sharp}{2} + V^\flat\ ,\ \frac{V^\sharp}{2} - V^\flat\right). \tag{2.33}$$

A **cooperative-competitive solution** *(or* **co-co solution***, in short) of the game described at (2.30) is defined as a pair of strategies* $(a^\sharp, b^\sharp)$ *together with a side payment* p *from Player B to Player A, such that*

$$\Phi^A(a^\sharp, b^\sharp) + p \;=\; \frac{V^\sharp}{2} + V^\flat, \qquad\qquad \Phi^B(a^\sharp, b^\sharp) - p \;=\; \frac{V^\sharp}{2} - V^\flat.$$

Here $V^\sharp$ *is maximum combined payoff defined at (2.31), while* $V^\flat$ *is the value of the zero-sum game with payoff* $\Phi^\flat$ *defined at (2.32).*

The concept of co-co solution models a situation where the players join forces, implement a strategy $(a^\sharp, b^\sharp)$ which achieves their maximum combined payoff. Then one of the two makes a side payment to the other, so that in the end the payoffs (2.33) are achieved.

Example 7. Consider the bi-matrix game described in Fig. 9. In this case we have $V^\sharp = 12$, $V^\flat = 2$. Observe that (a_2, b_2) is a saddle point for the corresponding zero-sum game.

In a co-co solution of the game, the players should receive the payoffs $V^\sharp + V^\flat \;=\; 8$ and $V^\sharp - V^\flat \;=\; 4$ respectively. A co-co solution is thus provided by the pair of strategies (a_1, b_2) together with the side payment $p = 5$ (from Player B to Player A). A second co-co solution is given by (a_2, b_1), with side payment $p = -1$ (i.e. with side payment 1 from A to B).

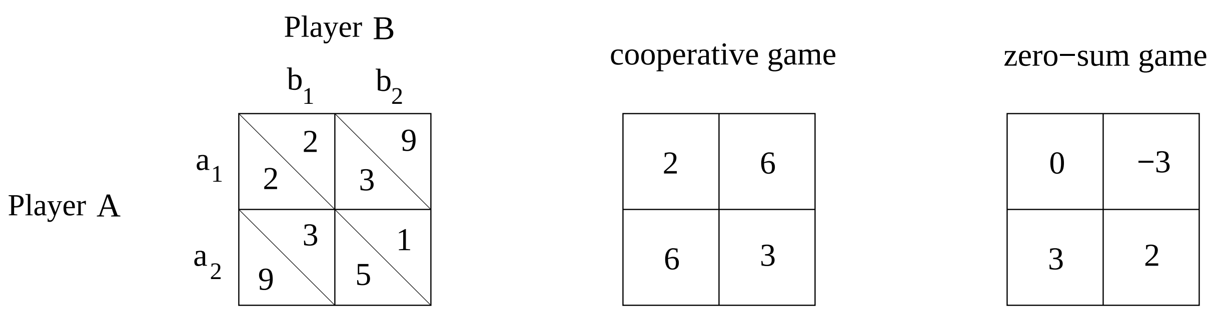

Figure 9: Left: the payoffs of a bi-matrix game, where each player has two options. This can be represented as the sum of a cooperative game where both players have exactly the same payoff, plus a zero-sum game. In the center is the matrix of payoffs $\Phi^\sharp$, on the right is the matrix of payoffs $\Phi^\flat$.

Example 8 (co-co solution of the prisoners' dilemma). For the game with payoff matrix given in Fig. 5, the co-co solution is (N, N). Since the game is symmetric, the corresponding zero-sum game has value $V^\flat = 0$ and no side payment is needed.

Problems

1. Consider a variation of the rock-paper-scissors game. Assume that player A has all three options, but player B can only choose rock or paper, i.e. the choice "scissors" is not available to him. Find a pair of randomized strategies yielding a Nash equilibrium, and compute the corresponding value of the game.

2. Consider a zero-sum game where the set of strategies for the two players are $A = B = [0,1]$, with payoff function $\Phi(a,b) = |a-b|$. Compute the upper and lower value of the game. Find a pair of randomized strategies yielding a Nash equilibrium. Compute the corresponding value of the game.

3. Let $x_1, x_2 \in I\!R$ be the strategies implemented by the two players, and let the corresponding payoffs be

$$\begin{aligned} \Phi^A(x_1,x_2) &= a_1x_1 + a_2x_2 - \left[a_{11}\frac{x_1^2}{2} + a_{12}x_1x_2 + a_{22}\frac{x_2^2}{2}\right], \\ \Phi^B(x_1,x_2) &= b_1x_1 + b_2x_2 - \left[b_{11}\frac{x_1^2}{2} + b_{12}x_1x_2 + b_{22}\frac{x_2^2}{2}\right], \end{aligned}$$

with $a_{11} > 0$, $b_{22} > 0$.

(i) Compute the best reply maps.

(ii) Compute a Nash equilibrium solution.

(iii) Compute a Stackelberg equilibrium solution.

(iv) For which values of the coefficients is the Stackelberg equilibrium better than the Nash equilibrium, for Player A?

3 Differential Games

From now on we consider games in continuous time. Let $x \in I\!R^N$ describe the state of the system, evolving in time according to the ODE

$$\dot{x}(t) = f(t,x,u_1,u_2) \qquad t \in [0,T], \tag{3.1}$$

with initial data

$$x(0) = x_0. \tag{3.2}$$

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

$$u_1 \in U_1, \qquad u_2 \in U_2, \tag{3.3}$$

for some given sets $U_1, U_2 \subseteq I\!R^m$.

For $i = 1,2$, the goal of the i-th player is to maximize his own payoff, namely

$$J_i(u_1,u_2) \doteq \psi_i(x(T)) - \int_0^T L_i(t,\, x(t),\, u_1(t),\, u_2(t))\,dt. \tag{3.4}$$

Here 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 vastly 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, and 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 $\mathcal{S}_i$ of strategies available to the i-th player will thus consist of all measurable functions $t \mapsto u_i(t)$ 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 $\mathcal{S}_i$ of strategies available to the i-th player will thus consist of all measurable functions $(t, x) \mapsto u_i(t, x)$ from $[0, T] \times \mathbb{R}^n$ into U_i.

CASE 3 (hierarchical play): Player 1 (the leader) announces his strategy in advance. This can be either open loop $u_1 = u_1^\clubsuit(t)$, or feedback $u_1 = u_1^\clubsuit(t, x)$. At this stage, the game yields an optimal control problem for Player 2 (the follower). Namely

$$\text{maximize:} \quad \psi_2(x(T)) - \int_0^T L_2\Big(t,\, x(t),\, u_1^\clubsuit(t, x(t)),\, u_2(t)\Big)\, dt\,, \tag{3.5}$$

subject to

$$\dot{x}(t) \;=\; f(t, x, u_1^\clubsuit(t, x)\,, u_2) \qquad x(0) = x_0\,, \qquad\qquad u_2(t) \in U_2\,. \tag{3.6}$$

Notice that in this case the knowledge of the initial point x_0 together with the evolution equation (3.6) 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 $u_1 = u_1^\clubsuit(t, x)$ such that the reply u_2 of the other player yields a payoff (for Player 1) as large as possible.

CASE 4 (delayed information): Assume that each player cannot observe the state $x(t)$, but gets information about the actions taken by the other player, with a time delay $\delta > 0$. In other words, assume that at each time $t > 0$ Player i gets to know the strategy $\{u_j(s)\,;\ \ s \in [0,\, t-\delta]\}$ adopted earlier by the other player.

This is a situation where cooperative agreements among the two players can be implemented, using strategies that trigger a punishment for the player who "cheats", deviating from the agreed course of action. For example, assume that the players agree in advance to adopt the controls $t \mapsto u_1^\heartsuit(t)$, $t \mapsto u_2^\heartsuit(t)$, yielding a Pareto optimum. However, assume that after time τ Player 2 changes his mind and adopts a different control, say

$$t \ \mapsto\ u_2^\spadesuit(t) \ \neq\ u_2^\heartsuit(t) \qquad\qquad t > \tau.$$

Here $u_2^\spadesuit$ can be a control that increases the payoff of Player 2 at the expense of Player 1. After time $t = \tau + \delta$, Player 1 discovers that he has been cheated. He can then punish his partner, choosing a new control $t \mapsto u_1^\spadesuit(t)$ yielding a very low payoff to Player 2.

Remark 4. At first sight, the threat of a punishment should induce both players to stick to their original agreement and implement the cooperative strategies $(u_1^\heartsuit,\, u_2^\heartsuit)$ during the entire time interval $t \in [0, T]$. However one should keep in mind that, by punishing Player 2 if he cheats, also Player 1 may have to reduce his own payoff as well. Since delivering a punishment can be very costly, in each given situation one should carefully evaluate whether the threat of punishment by one player to the other is credible or not.

4 Open loop strategies

In this section we consider solutions to the differential game (3.1)–(3.4), in the case where the strategies implemented by the players must be functions of time alone.

4.1 Open-loop Nash equilibrium solutions

Definition (open-loop Nash equilibrium). *A pair of control functions $t \mapsto (u_1^*(t),\, u_2^*(t))$ is a Nash equilibrium for the game (3.1)–(3.4) 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:

$$\textit{maximize:} \qquad J_1(u_1, u_2^*) \ = \ \psi_1(x(T)) - \int_0^T L_1\Big(t,\, x(t),\, u_1(t),\, u_2^*(t)\Big)\, dt\,. \tag{4.1}$$

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

$$x(0) = x_0 \in I\!R^N\,, \qquad \dot x(t) \ = \ f(t, x, u_1, u_2^*(t))\,, \qquad u_1(t) \in U_1 \qquad\qquad t \in [0, T]\,. \tag{4.2}$$

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

$$\textit{maximize:} \qquad J_2(u_1^*, u_2) \ = \ \psi_2(x(T)) - \int_0^T L_2\Big(t,\, x(t),\, u_1^*(t),\, u_2(t)\Big)\, dt\,. \tag{4.3}$$

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

$$x(0) = x_0 \in \mathbb{R}^N\,, \qquad \dot x(t) \;=\; f(t,x,u_1^*(t),u_2(t))\,, \qquad u_2(t)\in U_2 \qquad\qquad t\in[0,T]\,. \tag{4.4}$$

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,ψ_1,ψ_2,L_1,L_2 are continuously differentiable, necessary conditions for optimality are provided by the Pontryagin Maximum Principle, see Theorem A.7 in the Appendix.

Based on the PMP, we now describe a procedure for finding a pair of open-loop strategies $t\mapsto (u_1^*(t),\, u_2^*(t))$ yielding a Nash equilibrium. Toward this goal, we need to assume that a family of pointwise maximization problems can be uniquely solved. More precisely, we assume

(A2) *For every $(t,x)\in[0,T]\times\mathbb{R}^N$ and any two vectors $q_1,q_2\in\mathbb{R}^N$, there exists a unique pair $(u_1^\sharp,\, u_2^\sharp)\in U_1\times U_2$ such that*

$$u_1^\sharp \;=\; \operatorname*{argmax}_{\omega\in U_1}\Big\{q_1\cdot f(t,x,\omega,u_2^\sharp) - L_1(t,x,\omega,u_2^\sharp)\Big\}\,, \tag{4.5}$$

$$u_2^\sharp \;=\; \operatorname*{argmax}_{\omega\in U_2}\Big\{q_2\cdot f(t,\,x,\,u_1^\sharp,\omega) - L_2(t,x,u_1^\sharp,\,\omega)\Big\}\,. \tag{4.6}$$

The corresponding map will be denoted by

$$(t,x,q_1,q_2)\;\mapsto\;\Big(u_1^\sharp(t,x,q_1,q_2)\,,\;u_2^\sharp(t,x,q_1,q_2)\Big). \tag{4.7}$$

The assumption (A2) can be interpreted as follows. For any given $(t,x,q_1,q_2)\in[0,T]\times\mathbb{R}^N\times\mathbb{R}^N\times\mathbb{R}^N$, consider the "one-shot" game where the players choose strategies $u_i\in U_i$ in order to maximize their instantaneous payoffs

$$\Phi_i(u_1,u_2) \;=\; q_i\cdot f(t,x,u_1,u_2) - L_i(t,x,u_1,u_2) \qquad\qquad i=1,2\,. \tag{4.8}$$

According to (A2), 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 (4.7) is continuous. We now describe an important class of problems where this assumption is satisfied.

Lemma 2. *Assume that the dynamics and the running costs take the decoupled form*

$$f(t,x,u_1,u_2) \;=\; f_0(t,x) + M_1(t,x)u_1 + M_2(t,x)u_2\,, \tag{4.9}$$

$$L_i(t,x,u_1,u_2) \;=\; L_{i1}(t,x,u_1) + L_{i2}(t,x,u_2). \tag{4.10}$$

Assume that

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

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

(iii) The functions $u_1 \mapsto L_{11}(t, x, u_1)$ and $u_2 \mapsto L_{22}(t, x, u_2)$ 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 (A2) holds.

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

$$u_1^\sharp = \operatorname*{argmax}_{\omega\in U_1} \Big\{ q_1 \cdot M_1(t,x)\omega - L_{11}(t,x,\omega) \Big\}, \qquad u_2^\sharp = \operatorname*{argmax}_{\omega\in U_2} \Big\{ q_2 \cdot M_2(t,x)\omega - L_{22}(t,x,\omega) \Big\}. \tag{4.11}$$

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

Finding a Nash equilibrium using the PMP. Assume that (A2) 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 (see Theorem A.7 in the Appendix), one obtains the following set of necessary conditions.

$$\begin{cases} \dot x = f(t,\, x,\, u_1^\sharp,\, u_2^\sharp), \\ \dot q_1 = -\, q_1 \dfrac{\partial f}{\partial x}(t,\, x,\, u_1^\sharp,\, u_2^\sharp) + \dfrac{\partial L_1}{\partial x}(t,\, x,\, u_1^\sharp,\, u_2^\sharp)\,, \\ \dot q_2 = -\, q_2 \dfrac{\partial f}{\partial x}(t,\, x,\, u_1^\sharp,\, u_2^\sharp) + \dfrac{\partial L_2}{\partial x}(t,\, x,\, u_1^\sharp,\, u_2^\sharp)\,, \end{cases} \tag{4.12}$$

with initial and terminal conditions

$$\begin{cases} x(0) = x_0\,, \\ q_1(T) = \nabla\psi_1(x(T))\,, \\ q_2(T) = \nabla\psi_2(x(T))\,. \end{cases} \tag{4.13}$$

Notice that in (4.12) the variables $u_1^\sharp, u_2^\sharp$ are functions of (t, x, q_1, q_2), defined at (4.5)–(4.7).

One can use the above system in order to compute a Nash equilibrium solution to the differential game. Notice that (4.12) consists of three ODEs in $I\!R^N$. This needs to be solved with the mixed boundary data (4.13). 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 nonlinearity of the maps $u_1^\sharp, u_2^\sharp$ in (4.7), this makes the problem (4.12)-(4.13) hard to solve, in general.

As soon as a solution $t \mapsto (x(t),\ q_1(t),\ q_2(t))$ to the two-point boundary value problem (4.12)-(4.13) is found, the trajectory x^* and the controls u_1^*, u_2^* are determined by

$$x^*(t) = x(t)\,, \qquad u_1^*(t) \;=\; u_1^\sharp(t, x(t), q_1(t), q_2(t))\,, \qquad u_2^*(t) \;=\; u_2^\sharp(t, x(t), q_1(t), q_2(t))\,.$$

One should keep in mind that the Pontryagin maximum principle is only a necessary condition, not sufficient for optimality. In other words, any pair $t \mapsto (u_1^*(t), u_2^*(t))$ of open-loop strategies which is a Nash equilibrium must provide a solution to (4.12)-(4.13). On the other hand, being a solution of (4.12)-(4.13) does not guarantee that the pair (u_1^*, u_2^*) is a Nash equilibrium. A (very restrictive) setting where the PMP is actually sufficient for optimality is described in Theorem A.9 of the Appendix.

Example 9 (duopolistic competition). Two companies sell the same product, competing for market share. Let $x_1 = x(t) \in [0,1]$ be the market share of the first company at time t, so that $x_2 = 1 - x(t)$ is the market share of the second. Calling $u_i(t)$ be the advertising effort of firm $i \in \{1,2\}$ at time t, the Lanchester model is described by the dynamics

$$\dot x \;=\; (1-x)u_1 - x u_2\,, \qquad\qquad x(0) = x_0 \in [0,1]\,. \tag{4.14}$$

The i-th firm should plan its strategy $t \mapsto u_i(t)$ in order to maximize the total payoff

$$J_i \;\doteq\; \int_0^T \left[a_i\, x_i(t) - c_i\, \frac{u_i^2(t)}{2} \right] dt + S_i x_i(T)\,, \tag{4.15}$$

for suitable constants $a_i, c_i, S_i > 0$. Here the term $a_i x$ accounts for the earnings of the i-th company, proportional to its market share, while $c_i u_i^2/2$ is the advertising cost. The value attached by firm i to its terminal market share is described by $S_i x_i(T)$. A Nash equilibrium solution to this differential game, in terms of open-loop controls, can be found as follows.

STEP 1: the optimal controls are determined in terms of the adjoint variables:

$$\begin{cases} u_1^\sharp(x, q_1, q_2) = \displaystyle\operatorname*{argmax}_{\omega \ge 0} \left\{ q_1 \cdot (1-x)\omega - c_1\, \frac{\omega^2}{2} \right\} \;=\; (1-x)\, \frac{q_1}{c_1}\,, \\[2ex] u_2^\sharp(x, q_1, q_2) = \displaystyle\operatorname*{argmax}_{\omega \ge 0} \left\{ q_2 \cdot x\omega - c_2\, \frac{\omega^2}{2} \right\} \;=\; \frac{q_2}{c_2}\,. \end{cases} \tag{4.16}$$

STEP 2: the state $x(\cdot)$ and the adjoint variables $q_1(\cdot), q_2(\cdot)$ are determined by solving the boundary value problem

$$\begin{cases} \dot x \;=\; (1-x)u_1^\sharp + x u_2^\sharp \;=\; (1-x)^2\, \dfrac{q_1}{c_1} + x^2\, \dfrac{q_2}{c_2}\,, \\[2ex] \dot q_1 \;=\; -\, q_1\, (u_1^\sharp + u_2^\sharp) - a_1 \;=\; -\, q_1 \left[(1-x)\, \dfrac{q_1}{c_1} + x\, \dfrac{q_2}{c_2} \right] - a_1\,, \\[2ex] \dot q_2 \;=\; -\, q_2\, (u_1^\sharp + u_2^\sharp) - a_2 \;=\; -\, q_2 \left[(1-x)\, \dfrac{q_1}{c_1} + x\, \dfrac{q_2}{c_2} \right] - a_2\,, \end{cases} \tag{4.17}$$

with initial and terminal conditions

$$\begin{cases} x(0) &= x_0\,, \\ q_1(T) &= S_1\,, \\ q_2(T) &= S_2\,. \end{cases} \tag{4.18}$$

Example 10 (producer-consumer game with sticky price). Let $p(t)$ denote the price of a good at time t. We assume that this good can be produced by one of the players, at rate $u_1(t)$, and consumed by the other player at rate $u_2(t)$. In a very simplified model, the variation of the price in time can be described by the differential equation

$$\dot p \;=\; (u_2 - u_1)p\,, \tag{4.19}$$

Here the non-negative functions $t \mapsto u_1(t)$ and $t \mapsto u_2(t)$ represent the *controls* implemented by the two players. According to (4.19), the price increases when the consumption is larger than the production, and decreases otherwise.

Let the payoffs for the two players be described by

$$J^{prod} \;=\; \int_0^T \Big[p(t)u_2(t) - c(u_1(t))\Big]\,dt\,, \tag{4.20}$$

$$J^{cons} \;=\; \int_0^T \Big[\phi(u_2(t)) - p(t)u_2(t)\Big]\,dt\,. \tag{4.21}$$

The payoff for the producer is given by the profit generated by sales, minus the cost $c(u_1)$ of producing the good at rate u_1. The payoff for the consumer is measured by a utility function $\phi(u_2)$, minus the price payed to buy the good. For sake of definiteness, assume

$$c(s) = \frac{s^2}{2}\,, \qquad\qquad \phi(s) = 2\sqrt{s}\,. \tag{4.22}$$

A Nash equilibrium solution for this differential game, in terms of open-loop controls, is found as follows.

STEP 1: the optimal controls are determined in terms of the adjoint variables:

$$u_1^\sharp(x, q_1, q_2) \;=\; \operatorname*{argmax}_{\omega\geq 0}\;\left\{q_1\cdot(-\omega p) - \frac{\omega^2}{2}\right\} \;=\; -\,q_1 p\,,$$

$$u_2^\sharp(x, q_1, q_2) \;=\; \operatorname*{argmax}_{\omega\geq 0}\;\Big\{q_2\cdot(\omega p) + 2\sqrt{\omega} - p\omega\Big\} \;=\; \frac{1}{(1-q_2)^2\,p^2}\,.$$

Notice that here we are assuming $p > 0$, $q_1 \leq 0$, $q_2 < 1$.

STEP 2: the state $p(\cdot)$ and the adjoint variables $q_1(\cdot), q_2(\cdot)$ are determined by solving the boundary value problem

$$\begin{cases} \dot p \;=\; (u_2^\sharp - u_1^\sharp)p \;=\; \dfrac{1}{(q_2-1)^2\,p} + q_1 p^2\,, \\[2ex] \dot q_1 \;=\; -q_1(u_2^\sharp - u_1^\sharp) - u_2^\sharp \;=\; -\,q_1^2\,p - \dfrac{q_1+1}{(1-q_2)^2\,p^2}\,, \\[2ex] \dot q_2 \;=\; -q_2(u_2^\sharp - u_1^\sharp) + u_2^\sharp \;=\; -\,q_1 q_2 p + \dfrac{1}{(1-q_2)p}\,, \end{cases} \tag{4.23}$$

with initial and terminal conditions

$$\begin{cases} x(0) &= x_0\,, \\ q_1(T) &= 0\,, \\ q_2(T) &= 0\,. \end{cases} \tag{4.24}$$

4.2 Open-loop Stackelberg equilbrium solutions

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 $t \mapsto u_1(t)$, 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 $u_1^\clubsuit : [0,T] \mapsto U_1$ for the first player, we denote by $\mathcal{R}_2(u_1^\clubsuit)$ the set of **best replies** for the second player. More precisely, $\mathcal{R}_2(u_1^\clubsuit)$ is the set of all admissible control functions $u_2 : [0,T] \mapsto U_2$ for Player 2, which achieve the maximum payoff in connection with $u_1^\clubsuit$. Namely, they solve the optimal control problem

$$\text{maximize:} \quad \psi_2(x(T)) - \int_0^T L_2\Big(t,\, x(t),\, u_1^\clubsuit(t),\, u_2(t)\Big)\, dt\,, \tag{4.25}$$

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

$$\dot{x}(t) \;=\; f(t, x, u_1^\clubsuit(t)\,, u_2), \qquad\qquad x(0) = x_0\,, \qquad\qquad u_2(t) \in U_2\,. \tag{4.26}$$

In the following, given two control functions $u_1(\cdot)$ and $u_2(\cdot)$, we denote by $t \mapsto x(t, u_1, u_2)$ the solution of the Cauchy problem

$$\dot{x} \;=\; f(t, x, u_1(t), u_2(t))\,, \qquad\qquad x(0) = x_0\,.$$

Definition (open-loop Stackelberg equilibrium). *We say that a pair of control functions $t \mapsto (u_1^*(t),\, u_2^*(t))$ is a Stackelberg equilibrium for the game (3.1)–(3.4) within the class of open-loop strategies if the following holds.*

(i) $u_2^* \in \mathcal{R}(u_1^*)\,,$

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

$$\begin{aligned} &\psi_1(x(T, u_1, u_2)) - \int_0^T L_1\Big(t, x(t, u_1, u_2), u_1(t), u_2(t)\Big)\, dt \\ &\qquad \leq \; \psi_1(x(T, u_1^*, u_2^*)) - \int_0^T L_1\Big(t, x(t, u_1^*, u_2^*), u_1^*(t), u_2^*(t)\Big)\, dt\,. \end{aligned} \tag{4.27}$$

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 $u_1^*(\cdot)$ 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 $t \mapsto x^*(t)$ 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

$$\begin{cases} \dot{x}^*(t) \;=\; f(t, x^*(t), u_1^*(t), u_2^*(t))\,, \\[2ex] \dot{q}_2^*(t) \;=\; -\,q_2^* \cdot \dfrac{\partial f}{\partial x}(t, x^*(t), u_1^*(t), u_2^*(t)) + \dfrac{\partial L_2}{\partial x}(t, x^*(t), u_1^*(t), u_2^*(t))\,, \end{cases} \tag{4.28}$$

with boundary conditions

$$\begin{cases} x^*(0) \;=\; x_0 \\[2ex] q_2^*(T) \;=\; \nabla\psi_2(x^*(T))\,. \end{cases} \tag{4.29}$$

Moreover, the following optimality conditions hold:

$$u_2^*(t) \;\in\; \operatorname*{argmax}_{\omega\in U_2} \Big\{ q_2^*(t)\cdot f(t, x^*(t), u_1^*(t), \omega) - L_2(t, x^*(t), u_1^*(t), \omega) \Big\} \qquad \text{for a.e. } t \in [0,T]\,. \tag{4.30}$$

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

(A3) For each $(t, x, u_1, q_2) \;\in\; [0,T]\times I\!R^n \times U_1 \times I\!R^n$, there exists a unique optimal choice $u_2^\flat \in U_2$ for Player 2, namely

$$u_2^\flat(t, x, u_1, q_2) \;\doteq\; \operatorname*{argmax}_{\omega\in U_2} \Big\{ q_2 \cdot f(t, x, u_1, \omega) - L_2(t, x, u_1, \omega) \Big\}. \tag{4.31}$$

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 $(x, q_2) \in I\!R^n \times I\!R^n$.

$$\text{Maximize :} \qquad \psi_1(x(T)) - \int_0^T L_1\Big(t, x(t),\, u_1(t),\, u_2^\flat(t, x(t), u_1(t), q_2(t))\Big)\, dt \tag{4.32}$$

for the system on $I\!R^{2n}$ with dynamics

$$\begin{cases} \dot{x}(t) \;=\; f(t,\, x,\, u_1,\, u_2^\flat(t, x, u_1, q_2))\,, \\[2ex] \dot{q}_2(t) \;=\; -\,q_2 \cdot \dfrac{\partial f}{\partial x}(t,\, x,\, u_1,\, u_2^\flat(t, x, u_1, q_2)) + \dfrac{\partial L_2}{\partial x}(t,\, x,\, u_1,\, u_2^\flat(t, x, u_1, q_2))\,, \end{cases} \tag{4.33}$$

and with boundary conditions

$$x(0) \;=\; x_0, \qquad\qquad q_2(T) \;=\; \nabla\psi_2(x(T))\,. \tag{4.34}$$

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 $q_2 = \nabla\psi_2(x)$ valid at $t = T$. In order to apply the PMP, we need to assume that all functions in (4.32)–(4.34) are continuously differentiable w.r.t. the new state variables x, q_2. More precisely

(A4) *For every fixed $t \in [0, T]$ and $u_1 \in U_1$, the maps*

$$
\begin{aligned}
(x, q_2) &\mapsto \widetilde{L}_1(t, x, u_1, q_2) \doteq L_1\Big(t,\, x,\, u_1,\, u_2^\flat(t, x, u_1, q_2)\Big)\,, \\
(x, q_2) &\mapsto F(t, x, u_1, q_2) \doteq f(t,\, x,\, u_1,\, u_2^\flat(t, x, u_1, q_2))\,, \\
(x, q_2) &\mapsto G(t, x, u_1, q_2) \\
&\qquad \doteq -q_2 \cdot \frac{\partial f}{\partial x}(t,\, x,\, u_1,\, u_2^\flat(t, x, u_1, q_2)) + \frac{\partial L_2}{\partial x}(t,\, x,\, u_1,\, u_2^\flat(t, x, u_1, q_2))\,, \\
x &\mapsto \nabla\psi_2(x)\,,
\end{aligned}
$$

are continuously differentiable.

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

Theorem 5 (necessary conditions for an open-loop Stackelberg equilibrium). *Let the assumptions (A3)-(A4) hold. Let $t \mapsto (u_1^*(t),\, u_2^*(t))$ be open-loop strategies yielding a Stackelberg equilibrium for the differential game (3.1)–(3.4). Let $x^*(\cdot)$, $q_2^*(\cdot)$ be the corresponding trajectory and adjoint vector for Player 2, satisfying (4.28)–(4.30).*

Then there exists a constant $\lambda_0 \geq 0$ and two absolutely continuous adjoint vectors $\lambda_1(\cdot)$, $\lambda_2(\cdot)$ (not all equal to zero), satisfying the equations

$$
\begin{cases}
\dot{\lambda}_1 = \lambda_0 \dfrac{\partial \widetilde{L}_1}{\partial x} - \lambda_1 \dfrac{\partial F}{\partial x} - \lambda_2 \dfrac{\partial G}{\partial x}\,, \\[2ex]
\dot{\lambda}_2 = \lambda_0 \dfrac{\partial \widetilde{L}_1}{\partial q_2} - \lambda_1 \dfrac{\partial F}{\partial q_2} - \lambda_2 \dfrac{\partial G}{\partial q_2}\,,
\end{cases}
\tag{4.35}
$$

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

$$
\lambda_2(0) = 0\,, \qquad \lambda_1(T) = \lambda_0 \nabla\psi_1(x^*(T)) - \lambda_2(T)\, D^2\psi_2(x^*(T))\,. \tag{4.36}
$$

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

$$
u_1^*(t) = \underset{\omega \in U_1}{\operatorname{argmax}} \left\{ -\lambda_0 \widetilde{L}_1\Big(t, x^*(t), q_2^*(t), \omega\Big) + \lambda_1(t) \cdot F\Big(t, x^*(t), q_2^*(t), \omega\Big) + \lambda_2(t) \cdot G\Big(t,\, x^*(t),\, q_2^*(t),\, \omega\Big) \right\}. \tag{4.37}
$$

In the ODEs (4.35), it is understood that the right hand sides are computed at $(t, x^*(t), q_2^*(t), u_1^*(t))$. In (4.36), by $D^2\psi_2(x)$ we denote the Hessian matrix of second derivatives of ψ_2, at the point x.

The above result follows by applying Theorem A.8 in the Appendix the the optimal control problem (4.32)–(4.34). Observe that the initial data is constrained to the set

$$S_0 \;=\; \{(x, q_2) \in I\!R^{n+n}\,; \quad x = x_0\}\,.$$

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

$$\lambda_1(0) \;\in\; I\!R^n\,, \qquad\qquad \lambda_2(0) \;=\; 0.$$

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

$$S_T \;=\; \big\{(x, q_2) \in I\!R^{n+n}\,; \quad q_2 - \nabla\psi_2(x) = 0\big\}\;.$$

A vector $(\mathbf{v}_1,\, \mathbf{v}_2) \in I\!R^{2n}$ is tangent to the manifold S_T at the point (x, q_2) provided that

$$\mathbf{v}_2 \;=\; D_x^2\psi_2(x)\mathbf{v}_1\,.$$

Hence a vector $(\mathbf{n}_1,\, \mathbf{n}_2) \in I\!R^{2n}$ is normal to S_T provided that

$$\mathbf{n}_1 \;=\; -\,D_x^2\psi_2(x)\mathbf{n}_2\,.$$

Recalling that the terminal payoff is $\psi_1(x(T))$, from (9.29) we obtain the terminal condition

$$\lambda_1(T) \;=\; \lambda_0\,\nabla\psi_1(x^*(T)) - \lambda_2(T)\,D^2\psi_2(x(T))\,,$$

for some constant $\lambda_0 \geq 0$.

Example 11 (economic growth). Let $x(t)$ describe the total wealth of capitalists in a country, at time t. Assume that this quantity evolves according to

$$\dot{x} \;=\; a\,x - u_1 x - u_2\,, \qquad\qquad x(0) = x_0\,, \qquad t \in [0, T]\,. \tag{4.38}$$

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

$$J_1 \;=\; bx(T) + \int_0^T \phi_1\Big(u_1(t)x(t)\Big)\,dt\,, \tag{4.39}$$

$$J_2 \;=\; x(T) + \int_0^T \phi_2\Big(u_2(t)\Big)\,dt\,. \tag{4.40}$$

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 $u_1(\cdot)$ as a function of time, and the capitalists are the followers. For this example, the functions considered in (A3)-(A4) take the form

$$u_2^\flat(x, u_1, q_2) \;=\; \operatorname*{argmax}_{\omega \geq 0}\Big\{-\,q_2\omega + k_2\ln\omega\Big\} \;=\; \frac{k_2}{q_2}\,,$$

$$\begin{aligned}
\widetilde{L}_1(x, q_2, u_1) &= \phi_1(u_1 x) = k_1 \ln(u_1 x)\,,\\
F(x, q_2, u_1) &= a\,x - u_1 x - \frac{k_2}{q_2}\,,\\
G(x, q_2, u_1) &= -\,q_2(a - u_1)\,.
\end{aligned}$$

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

$$\text{maximize:} \qquad bx(T) + \int_0^T k_1 \ \ln\Big(u_1(t)x(t)\Big)\,dt\,, \tag{4.41}$$

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

$$\begin{cases}
\dot{x} = ax - u_1 x - \dfrac{k_2}{q_2}\,,\\[2ex]
\dot{q}_2 = -\,q_2(a - u_1)\,,
\end{cases} \tag{4.42}$$

and boundary conditions

$$x(0) = x_0\,, \qquad\qquad q_2(T) = 1\,. \tag{4.43}$$

By the Pontryagin maximum principle (Theorem A.8 in the Appendix), an optimal control can be found as follows.

STEP 1: For any constants $\lambda_0 \geq 0$, λ_1, λ_2, compute the optimal feedback control

$$u_1^\sharp(x, q_2, \lambda_0, \lambda_1, \lambda_2) \doteq \operatorname*{argmax}_{\omega \geq 0} \Big\{\lambda_1\,(-\omega x) + \lambda_2 q_2 \omega + \lambda_0\, k_1 \ln(\omega x)\Big\} = \frac{\lambda_0 k_1}{\lambda_1 x - \lambda_2 q_2}\,.$$

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

$$\begin{cases}
\dot{x} = (a - u_1^\sharp)x - \dfrac{k_2}{q_2} = \left(a - \dfrac{\lambda_0 k_1}{\lambda_1 x - \lambda_2 q_2}\right) x - \dfrac{k_2}{q_2}\,,\\[2ex]
\dot{q}_2 = -q_2(a - u_1^\sharp) = \left(\dfrac{\lambda_0 k_1}{\lambda_1 x - \lambda_2 q_2} - a\right) q_2\,,\\[2ex]
\dot{\lambda}_1 = -\lambda_0 \dfrac{\kappa_1}{x} - \lambda_1(a - u_1^\sharp) = -\,\lambda_0 \dfrac{\kappa_1}{x} + \lambda_1 \left(\dfrac{\lambda_0 k_1}{\lambda_1 x - \lambda_2 q_2} - a\right),\\[2ex]
\dot{\lambda}_2 = -\lambda_1 \dfrac{k_2}{q_2} + \lambda_2(a - u_1^\sharp) = -\,\lambda_1 \dfrac{k_2}{q_2} + \lambda_2 \left(a - \dfrac{\lambda_0 k_1}{\lambda_1 x - \lambda_2 q_2}\right),
\end{cases}$$

with initial and terminal conditions (see figure 10)

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

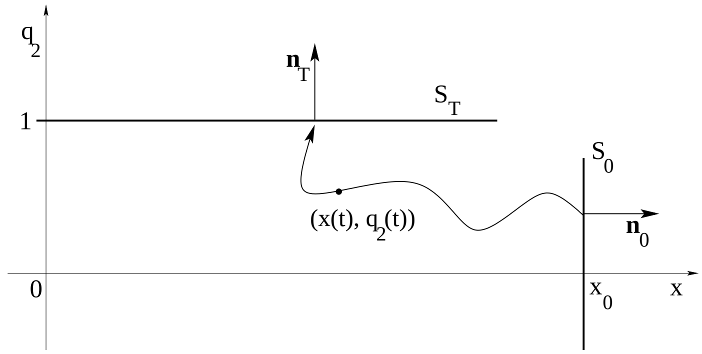

Figure 10: The initial and terminal constraints for the optimal control problem (4.41)–(4.43). According to (9.29), at time $t = 0$ one has $(\lambda_1(0),\, \lambda_2(0)) = \mathbf{n}_0$ for some vector $\mathbf{n}_0$ perpendicular to the manifold $S_0 = \{(x, q_2)\,;\;\; x = x_0\}$. Hence $\lambda_2(0) = 0$ while $\lambda_1(0)$ can be arbitrary. On the other hand, at time $t = T$ one has $(\lambda_1(T),\, \lambda_2(T)) = \lambda_0(b, 1) + \mathbf{n}_T$ for some vector $\mathbf{n}_T$ perpendicular to the manifold $S_T = \{(x, q_2)\,;\;\; q_2 = 1\}$. Hence $\lambda_1(T) = \lambda_0 b$ while $\lambda_2(T)$ can be arbitrary.

5 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 in (3.2). 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 (3.1), (3.4), simultaneously for any choice of initial data

$$x(\tau) \;=\; y, \tag{5.1}$$

with $\tau \in [0, T]$, $y \in I\!R^N$.

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

$$\max_u \;\left\{ \psi(x(T)) - \int_\tau^T L\Big(t,\, x,\, u\Big)\, dt \right\}, \tag{5.2}$$

with dynamics

$$\dot{x} \;=\; f(t, x, u)\,, \qquad\qquad u(t) \in U\,, \tag{5.3}$$

if, for every initial data $(\tau, y) \in [0, T] \times I\!R^N$, every Carathéodory solution of the Cauchy problem

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

is optimal, i.e. it achieves the maximum payoff in (5.2).

Definition (feedback Nash equilibrium). *A pair of control functions $(t, x) \mapsto (u_1^*(t, x),\, u_2^*(t, x))$ is a Nash equilibrium for the game (3.1), (3.3), (3.4) within the class of feedback strategies if the following holds.*

(i) The control $(t, x) \mapsto u_1^*(t, x)$ provides an optimal feedback in connection with the optimal

control problem for Player 1:

$$\max_{u_1} \left\{ \psi_1(x(T)) - \int_0^T L_1\Big(t,\, x(t),\, u_1\,,\, u_2^*(t,x(t))\Big)\,dt \right\}, \tag{5.4}$$

for the system with dynamics

$$\dot{x}(t) \;=\; f(t,x,u_1,u_2^*(t,x))\,, \qquad u_1(t) \in U_1\,. \tag{5.5}$$

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

$$\max_{u_2} \left\{ \psi_2(x(T)) - \int_0^T L_2\Big(t,\, x(t),\, u_1^*(t,x(t)),\, u_2\Big)\,dt \right\}, \tag{5.6}$$

for the system with dynamics

$$\dot{x}(t) \;=\; f(t,x,u_1^*(t,x),u_2)\,, \qquad u_2 \in U_2\,. \tag{5.7}$$

5.1 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 $(\tau, y) \in [0,T] \times I\!R^N$, call $t \mapsto x^*(t;\tau,y)$ the solution of

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

We here assume that all these solutions are well defined. This is clearly true if the feedback controls u_1^*, u_2^* are Lipschitz continuous w.r.t. the variable x, but it is a nontrivial assumption in general.

We can then define the corresponding value functions V_1, V_2 as

$$V_i(\tau,y) \;=\; \psi_i(x^*(T)) - \int_\tau^T L_i\Big(t, x^*(t),\, u_1^*(t,x^*(t)),\, u_2^*(t,x^*(t))\Big)\,dt\,,$$

where $x^*(t) \doteq x^*(t,\tau,y)$. 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 (A2) hold. On a region where V_1, V_2 are $\mathcal{C}^1$, by the dynamic programming principle (see Theorem A.10 in the Apendix) they satisfy the system of Hamilton-Jacobi PDEs

$$\begin{cases} V_{1,t} + \nabla V_1 \cdot f(t,x,u_1^\sharp,u_2^\sharp) \;=\; L_1(t,x,u_1^\sharp,u_2^\sharp)\,, \\[2ex] V_{2,t} + \nabla V_2 \cdot f(t,x,u_1^\sharp,u_2^\sharp) \;=\; L_2(t,x,u_1^\sharp,u_2^\sharp)\,. \end{cases} \tag{5.8}$$

This system is closed by the equations

$$u_i^\sharp \;=\; u_i^\sharp(t,x,\,\nabla V_1,\,\nabla V_2) \qquad\qquad i = 1,2, \tag{5.9}$$

introduced at (4.7), and complemented by the terminal conditions

$$V_1(T,x) \;=\; \psi_1(x)\,, \qquad\qquad V_2(T,x) \;=\; \psi_2(x)\,. \tag{5.10}$$

Because of the nonlinearity of the functions $(t,x,q^1,q^2)\mapsto u_i^\sharp(t,x,q^1,q^2)$, the system (5.8) 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 (5.8), and let

$$V^\varepsilon(t,x) \;=\; V(t,x)+\varepsilon\, Z(t,x)+o(\varepsilon) \tag{5.11}$$

describe a small perturbation. Here the Landau symbol $o(\varepsilon)$ denotes a higher order infinitesimal, as $\varepsilon\to 0$. Assuming that V^ε is also a solution, we can insert (5.11) in the equation (5.8) and compute a linearized equation satisfied by the first order perturbation $Z=(Z_1,Z_2)$. Writing $f=(f_1,\ldots,f_n)$, $q_1=(q_{11},\ldots q_{1n})$, $q_2=(q_{21},\ldots q_{2n})$, we find

$$Z_{i,t}+\sum_{\alpha=1}^n f_\alpha\, Z_{i,x_\alpha}+\sum_{\alpha=1}^n\sum_{j=1}^2\left(\nabla V_i\cdot\frac{\partial f}{\partial u_j}-\frac{\partial L_i}{\partial u_j}\right)\left(\frac{\partial u_j^\sharp}{\partial q_{1\alpha}}\,Z_{1,x_\alpha}+\frac{\partial u_j^\sharp}{\partial q_{2\alpha}}\,Z_{2,x_\alpha}\right) \;=\; 0 \qquad i=1,2\,. \tag{5.12}$$

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

$$\nabla V_1\cdot\frac{\partial f}{\partial u_1}-\frac{\partial L_1}{\partial u_1} \;=\; 0\,, \qquad\qquad \nabla V_2\cdot\frac{\partial f}{\partial u_2}-\frac{\partial L_2}{\partial u_2} \;=\; 0. \tag{5.13}$$

Therefore, these terms drop off from the right hand sides of (5.12). In matrix notation, this homogeneous linear system can be written as

$$\begin{pmatrix} Z_{1,t}\\ Z_{2,t}\end{pmatrix}+\sum_{\alpha=1}^n A^\alpha\begin{pmatrix} Z_{1,x_\alpha}\\ Z_{2,x_\alpha}\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}, \tag{5.14}$$

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

$$A^\alpha \;=\; \begin{pmatrix} f_\alpha+\left(\nabla V_1\cdot\dfrac{\partial f}{\partial u_2}-\dfrac{\partial L_1}{\partial u_2}\right)\dfrac{\partial u_2^\sharp}{\partial q_{1\alpha}} & \left(\nabla V_1\cdot\dfrac{\partial f}{\partial u_2}-\dfrac{\partial L_1}{\partial u_2}\right)\dfrac{\partial u_2^\sharp}{\partial q_{2\alpha}} \\[3ex] \left(\nabla V_2\cdot\dfrac{\partial f}{\partial u_1}-\dfrac{\partial L_2}{\partial u_1}\right)\dfrac{\partial u_1^\sharp}{\partial q_{1\alpha}} & f_\alpha+\left(\nabla V_2\cdot\dfrac{\partial f}{\partial u_1}-\dfrac{\partial L_2}{\partial u_1}\right)\dfrac{\partial u_1^\sharp}{\partial q_{2\alpha}}\end{pmatrix}. \tag{5.15}$$

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

According to Theorem A.13, a necessary condition in order that the system (5.14) be hyperbolic (and hence that the linear Cauchy be well posed), is that for all $\xi\in I\!R^n$ the matrix

$$A(\xi) \;=\; \sum_\alpha A^\alpha\xi_\alpha \tag{5.16}$$

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.

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

In this case the function $u_1^\sharp$ in (4.5) does not depend on q_2, and similarly the function $u_2^\sharp$ in (4.6) does not depend on q_1. The 2×2 matrix $A(\xi)$ thus takes the simpler form

$$A(\xi) \;=\; \sum_{\alpha=1}^n \begin{pmatrix} f_\alpha \xi_\alpha & \left(\nabla V_1\cdot \dfrac{\partial f}{\partial u_2}-\dfrac{\partial L_1}{\partial u_2}\right)\dfrac{\partial u_2^\sharp}{\partial q_{2\alpha}}\,\xi_\alpha \\ \left(\nabla V_2\cdot \dfrac{\partial f}{\partial u_1}-\dfrac{\partial L_2}{\partial u_1}\right)\dfrac{\partial u_1^\sharp}{\partial q_{1\alpha}}\,\xi_\alpha & f_\alpha\,\xi_\alpha \end{pmatrix}.$$

Consider the two vectors

$$\mathbf{v} \;=\; (\mathbf{v}_1,\ldots,\mathbf{v}_n)\,, \qquad\qquad \mathbf{v}_\alpha \;\doteq\; \left(\nabla V_1\cdot \frac{\partial f}{\partial u_2}-\frac{\partial L_1}{\partial u_2}\right)\frac{\partial u_2^\sharp}{\partial q_{2\alpha}}\,, \tag{5.17}$$

$$\mathbf{w} \;=\; (\mathbf{w}_1,\ldots,\mathbf{w}_n)\,, \qquad\qquad \mathbf{w}_\alpha \doteq \left(\nabla V_2\cdot \frac{\partial f}{\partial u_1}-\frac{\partial L_2}{\partial u_1}\right)\frac{\partial u_1^\sharp}{\partial q_{1\alpha}}\,. \tag{5.18}$$

Observe that the matrix $A(\xi)$ in (5.16) has real eigenvalues if and only if the two inner products satisfy

$$(\mathbf{v}\cdot\xi)\,(\mathbf{w}\cdot\xi) \;\geq\; 0\,. \tag{5.19}$$

The condition (5.19) is satisfied for all $\xi\in I\!\!R^n$ if and only if the two vectors $\mathbf{v},\mathbf{w}$ 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 $a\mathbf{u}=b\mathbf{w}$.

In any dimension $n\geq 2$, this condition generically fails. Indeed, if $\mathbf{v},\mathbf{w}$ are linearly independent, we can find a vector of the form $\xi=\mathbf{v}-\theta\mathbf{w}$ which is perpendicular to $\mathbf{v}+\mathbf{w}$, so that (5.19) fails. Hence the system (5.8) is NOT hyperbolic, and the linearized Cauchy problem is ill-posed, both forward and backward in time.

Going back to the general case (5.15), 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

$$\mathbf{z} \;=\; (\mathbf{z}_1,\ldots,\mathbf{z}_n)\,, \qquad \mathbf{z}_\alpha \;=\; \left(\nabla V_1\cdot \frac{\partial f}{\partial u_2}-\frac{\partial L_1}{\partial u_2}\right)\frac{\partial u_2^\sharp}{\partial q_{1\alpha}}-\left(\nabla V_2\cdot \frac{\partial f}{\partial u_1}-\frac{\partial L_2}{\partial u_1}\right)\frac{\partial u_1^\sharp}{\partial q_{2\alpha}}\,. \tag{5.20}$$

This holds if and only if

$$(\mathbf{z}\cdot\xi)^2+4(\mathbf{v}\cdot\xi)(\mathbf{w}\cdot\xi)\;\geq\;0 \qquad\qquad \text{for all}\;\; \xi\in I\!\!R^n. \tag{5.21}$$

In space dimension $n\geq 3$, the condition (5.21) generically fails. Indeed, assume that the vectors $\mathbf{v},\mathbf{w},\mathbf{z}$ are linearly independent. Then we can find a nonzero vector

$$\xi\;\in\;\{\mathbf{z},\,\mathbf{v}+\mathbf{w}\}^\perp\cap\mathrm{span}\{\mathbf{v},\mathbf{w},\mathbf{z}\}\,.$$

With this choice, the quantity in (5.21) is strictly negative.

In space dimension $n = 2$, however, one may find situations where

$$\min_{\xi \in \mathbb{R}^2,\ |\xi|=1} \left\{ (\mathbf{z}\cdot\xi)^2 + 4(\mathbf{v}\cdot\xi)(\mathbf{w}\cdot\xi) \right\} \ > \ 0\,.$$

For example, if the vectors in (5.17), (5.18), (5.20) happen to be

$$\mathbf{v} = (1,0)\,, \qquad \mathbf{w} = (1,1)\,, \qquad \mathbf{z} = (0,2)\,,$$

then the system (5.14)-(5.15), in two space dimensions, would be locally hyperbolic. Indeed, for any $\xi = (\xi_1, \xi_2)$, one has

$$(\mathbf{z}\cdot\xi)^2 + 4(\mathbf{v}\cdot\xi)(\mathbf{w}\cdot\xi) \ = \ 4\xi_1^2 + 4\xi_1(\xi_1+\xi_2) \ = \ 3(\xi_1+\xi_2)^2 + (\xi_1-\xi_2)^2 \ \geq \ 0.$$

Remark 5. 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 (5.15) should be computed only at points (t, x, q_1, q_2) where $q_2 = \nabla V_2 = -\nabla V_1 = -q_1$. By (5.13), this yields

$$\nabla V_2 \cdot \frac{\partial f}{\partial u_1} - \frac{\partial L_2}{\partial u_1} \ = \ -\Big(\nabla V_1 \cdot \frac{\partial f}{\partial u_1} - \frac{\partial L_1}{\partial u_1}\Big) \ = \ 0\,,$$

$$\nabla V_1 \cdot \frac{\partial f}{\partial u_2} - \frac{\partial L_1}{\partial u_2} \ = \ -\Big(\nabla V_2 \cdot \frac{\partial f}{\partial u_2} - \frac{\partial L_2}{\partial u_2}\Big) \ = \ 0\,.$$

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 (5.8)–(5.10) 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 noncooperative games, can be found in [8]. This result is obtained differentiating the equations (5.8) w.r.t. the space variable x. This yields a nonlinear system of conservation laws for the variables $q_1 = V_{1,x}$ and $q_2 = V_{2,x}$. 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.

5.2 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

$$\dot{x} \ = \ A(t)x + B_1(t)u_1 + B_2(t)u_2\,, \tag{5.22}$$

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

$$J_i \;=\; \psi_i(x(T)) - \int_0^T L_i(t, x(t), u_1(t), u_2(t))\,dt\,, \tag{5.23}$$

$$\psi_i(x) \;=\; \frac{1}{2}x^\dagger \overline{M}_i x\,, \tag{5.24}$$

$$L_i(t, x, u_1, u_2) \;=\; \frac{|u_i|^2}{2} + \frac{1}{2}x^\dagger P_i(t)x + \sum_{j=1,2} x^\dagger Q_{ij}(t)u_j\,, \tag{5.25}$$

where the superscript † denotes transposition. Here $x \in I\!R^n$, $u_1 \in I\!R^{m_1}$, $u_2 \in I\!R^{m_2}$ are column vectors, A is an $n \times n$ matrix, $\overline{M}_i, P_i$ are $n \times n$ symmetric matrices, Q_{ij} and B_j are $n \times m_j$ matrices.

In this model, it is important that the controls u_1, u_2 range over the entire spaces $I\!R^{m_1}, I\!R^{m_2}$, without being restricted to a compact subset. Notice that the assumption (A2) certainly holds: the functions $u_i^\sharp$ in (4.5)-(4.6) are explicitly computed as

$$u_i^\sharp(t, x, q_i) \;=\; \arg\max_{\omega\in I\!R^{m_i}} \left\{ q_i B_i(t)\omega - \frac{|\omega|^2}{2} - x^\dagger Q_{ii}(t)\omega \right\} \;=\; \Big(q_i B_i(t) - x^\dagger Q_{ii}(t)\Big)^\dagger \qquad i = 1, 2\,. \tag{5.26}$$

Even if the backward Cauchy problem (5.8)–(5.10) 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

$$V_i(t, x) \;=\; \frac{1}{2}\,x^\dagger M_i(t)x\,. \tag{5.27}$$

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

$$\nabla V_i(t, x) \;=\; x^\dagger M_i(t)\,, \qquad\qquad V_{i,t}(t, x) \;=\; \frac{1}{2}\,x^\dagger \dot{M}_i(t)x\,, \tag{5.28}$$

$$u_i^\sharp(t, x, \nabla V_i(t, x)) \;=\; \Big(x^\dagger M_i(t)B_i(t) - x^\dagger Q_{ii}(t)\Big)^\dagger. \tag{5.29}$$

By (5.28) and (5.29), the functions V_i in (5.27) solve the system

$$V_{i,t} \;=\; L_i - \nabla V_i \cdot f \qquad\qquad i = 1, 2,$$

if and only if the following relations are satisfied

$$\begin{aligned}
&\frac{1}{2}x^\dagger \dot{M}_i x \\
&= \left[\frac{1}{2}\Big(x^\dagger M_i B_i - x^\dagger Q_{ii}\Big)\Big(x^\dagger M_i B_i - x^\dagger Q_{ii}\Big)^\dagger + \frac{1}{2}x^\dagger P_i x + \sum_{j=1,2} x^\dagger Q_{ij}\Big(x^\dagger M_i B_i - x^\dagger Q_{ii}\Big)^\dagger\right] \\
&\quad - x^\dagger M_i\left(Ax + \sum_{j=1,2} B_j\Big(x^\dagger M_j B_j - x^\dagger Q_{jj}\Big)^\dagger\right).
\end{aligned} \tag{5.30}$$

Notice that both sides of (5.30) are homogeneous quadratic polynomials w.r.t. the variable $x = (x_1, \ldots, x_n)$. The equality holds for every $x \in I\!R^n$ if and only if the following identity between $n \times n$ symmetric matrices is satisfied:

$$\begin{aligned} \frac{1}{2}\dot{M}_i &= \frac{1}{2}(M_iB_i - Q_{ii})(M_iB_i - Q_{ii})^\dagger + \frac{1}{2}P_i + \frac{1}{2}\sum_{j=1,2}\Big[Q_{ij}(M_iB_i - Q_{ii})^\dagger + (M_iB_i - Q_{ii})Q_{ij}^\dagger\Big] \\ &\quad -\frac{1}{2}(M_iA + A^\dagger M_i) - \frac{1}{2}\sum_{j=1,2}\Big[M_iB_j(M_jB_j - Q_{jj})^\dagger + (M_jB_j - Q_{jj})B_j^\dagger M_i\Big]. \end{aligned} \tag{5.31}$$

The equations (5.31) 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

$$M_1(T) = \overline{M}_1, \qquad M_2(T) = \overline{M}_2. \tag{5.32}$$

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 (5.31)-(5.32) has as solution on the entire interval $[0, T]$, then the formulas (5.28)-(5.29) yield the the optimal feedback controls $u_i^*(t, x) = u_i^\sharp(t, x, \nabla V_i(t, x))$.

Remark 6. The above approach can be applied to a more general class of non-homogeneous linear-quadratic games, with dynamics

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

and payoff functions (5.23), where

$$\psi_i(x) = \frac{1}{2}x^\dagger\overline{M}_i x + \bar{\mathbf{a}}_i \cdot x + \bar{e},$$

$$L_i(t, x, u_1, u_2) = \frac{1}{2}u_i^\dagger R_i(t)u_i + \frac{1}{2}x^\dagger P_i(t)x + \sum_{j=1,2} x^\dagger Q_{ij}(t)u_j + \sum_{j=1,2} S_{ij}(t)u_j + \mathbf{b}_i(t)\cdot x.$$

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:

$$V_i(t, x) = x^\dagger M_i(t)x + \mathbf{a}_i(t)\cdot x + e(t) \qquad i = 1, 2.$$

6 Time discretizations

Based on the theory of static games, a natural approach to dynamic games is to discretize time and approximate a dynamic game by a sequence of "one shot" games. In this section we discuss this approach. We recall that there are several different concepts of "solution" to a one-shot game:

- Pareto optimum

• Nash non-cooperative equilibrium

• Stackelberg equilibrium

• Co-co (cooperative-competitive) solution, with side payments.

Each concept leads to a different notion of solution, for a dynamic game in continuous time.

Consider again the differential game with dynamics (3.1) and payoff functionals (3.4). Given an integer $N \geq 1$, we partition the time interval $[0, T]$ into N equal subintervals, by setting

$$h \doteq \frac{T}{N}, \qquad \tau_j = j\,h = \frac{jT}{N}, \qquad j = 1, 2, \ldots, N. \tag{6.1}$$

We now consider a sequence of N "one-shot" games, and value functions $V_1(\tau_j, \cdot)$, $V_2(\tau_j, \cdot)$, defined as follows.

For $j = N$, set

$$V_i(\tau_N, x) \doteq \psi_i(x) \qquad i = 1, 2. \tag{6.2}$$

For any given state $x \in I\!R^n$, consider the one-shot game with payoff functions

$$\Phi_i(\tau_N, x, \omega_1, \omega_2) \doteq \psi_i\Big(x + h\, f(\tau_N, x, \omega_1, \omega_2)\Big) - h\, L_i(\tau_N, x, \omega_1, \omega_2). \tag{6.3}$$

Here the controls for the two players are

$$\omega_1 \in U_1, \qquad \omega_2 \in U_2. \tag{6.4}$$

Assume that this game has a unique solution, corresponding to the controls

$$\omega_1 \doteq u_1(\tau_N, x), \qquad \omega_2 \doteq u_2(\tau_N, x). \tag{6.5}$$

For $i = 1, 2$, let

$$V_i(\tau_{N-1}, x) \doteq \Phi_i\Big(\tau_N,\, x,\, u_1(\tau_N, x),\, u_2(\tau_N, x)\Big)$$

be the payoffs achieved by the two players, for this solution.

By backward induction, assume that the value functions V_1, V_2 have already been determined, for $t = \tau_j, \tau_{j+1}, \ldots, \tau_N = T$. We then consider the one-shot game on the subinterval $[\tau_{j-1}, \tau_j]$, with payoffs

$$\Phi_i(\tau_j, x, \omega_1, \omega_2)) \doteq V_i\Big(\tau_j\,,\ x + h\, f(\tau_j, x, \omega_1, \omega_2)\Big) - h\, L_i(\tau_j,\, x,\, \omega_1, \omega_2). \tag{6.6}$$

Assume that this game has a unique solution, corresponding to controls

$$\omega_1 \doteq u_1(\tau_j, x), \qquad \omega_2 \doteq u_2(\tau_j, x). \tag{6.7}$$

For $i = 1, 2$, let

$$V_i(\tau_{j-1}, x) \doteq \Phi_i\Big(\tau_j,\, x,\, u_1(\tau_j, x),\, u_2(\tau_j, x)\Big)$$

be the payoffs achieved by the two players, for this solution.

Continuing this backward induction procedure, we obtain value functions $V_1(\tau_j, x)$, $V_2(\tau_j, x)$ and controls $u_1(\tau_j, x)$, $u_2(\tau_j, x)$, defined for all $j = 0, 1, \ldots, N$ and $x \in I\!R^n$. For convenience, we extend these functions to the entire domain $[0, T] \times I\!R^n$, by setting

$$V_i^{(N)}(t,x) = V_i^{(N)}(\tau_j, x)\,, \qquad u_i^{(N)}(t,x) = u_i^{(N)}(\tau_j, x) \qquad \text{for all } \ t \in [\tau_j,\, \tau_{j+1}[\,.$$

Notice that we now inserted the superscript $^{(N)}$, to remind that these functions are obtained via a partition of $[0, T]$ into N subintervals.

To study differential games in continuous time, we now take the limit as $N \to \infty$, so that the time step $h = T/N \to 0$. Two natural questions arise.

(Q1) Letting $N \to \infty$, study whether the following limits exist:

$$V_i^{(N)}(t,x) \to V_i^*(t,x)\,, \qquad u_i^{(N)}(t,x) \to u_i^*(t,x)\,. \tag{6.8}$$

(Q2) Assuming that the limits exist, derive a system of PDEs satisfied by the value functions V_1^*, V_2^*.

In the following we shall content ourselves with a formal analysis, deriving a system of Hamilton-Jacobi PDEs for the value functions.

Fix an integer $N \geq 1$ and consider the time step $h = T/N$. Assume that at a given time $t = \tau_j$ the value functions $x \mapsto V_i^{(N)}(t, x)$, $i = 1, 2$, are continuously differentiable. According to (6.6), to compute the values $V_i^{(N)}(t-h,\, x)$ one has to solve the one-shot game with payoffs

$$\begin{aligned}\Phi_i(t,x,\omega_1,\omega_2) &= V_i^{(N)}\Big(t,\ x + h\, f(t,x,\omega_1,\omega_2)\Big) - h\, L_i(t,\, x,\, \omega_1,\, \omega_2)\\ &= V_i^{(N)}(t,x) + h\, \nabla V_i^{(N)}(t,x)\cdot f(t,x,\omega_1,\omega_2) - h\, L_i(t,\, x,\, \omega_1,\, \omega_2) + o(h)\,.\end{aligned} \tag{6.9}$$

Since $V_i^{(N)}(t, x)$ does not depend on ω_1, ω_2, neglecting higher order terms $o(h)$, the optimal strategies for the two players are the same as in the "instantaneous game" with payoffs

$$\Psi_i(\omega_1,\omega_2) \ = \ q_i \cdot f(t,x,\omega_1,\omega_2) - L_i(t,\, x,\, \omega_1,\, \omega_2) \qquad i = 1,2\,, \tag{6.10}$$

with $q_i = \nabla V_i(t, x)$.

Assume that, for every $(t, x) \in [0, T] \times I\!R^n$ and every couple of vectors $(q_1, q_2) \in I\!R^n \times I\!R^n$, we are given a unique solution to the game (6.10), say

$$\omega_1 \ = \ u_1^\sharp(t,x,q_1,q_2), \qquad \omega_2 \ = \ u_2^\sharp(t,x,q_1,q_2), \tag{6.11}$$

continuously depending on the parameters t, x, q_1, q_2. Of course, different concepts of solution (Pareto, Nash, Stackelberg) will lead to different functions $u_1^\sharp, u_2^\sharp$.

Denoting by $o(1)$ a quantity that approaches zero as $h \to 0$, we can now write

$$\frac{V_i^{(N)}(t-h,x) - V_i^{(N)}(t,x)}{h} \ = \ \nabla V_i^{(N)}(t,x)\cdot f(t,x,u_1^\sharp,u_2^\sharp) - \ L_i(t,\, x,\, u_1^\sharp,\, u_2^\sharp) + o(1)\,. \tag{6.12}$$

Letting $N \to \infty$ and assuming the convergence

$$V_i^{(N)}(t,x) \to V^*(t,x), \qquad \nabla V_i^{(N)}(t,x) \ \to \ \nabla V_i^*(t,x)\,,$$

$$\frac{V_i^{(N)}(t,x) - V_i^{(N)}(t-h,\,x)}{h} \ \to \ V_{i,t}^*(t,x)\,,$$

from (6.12) we conclude

$$\begin{cases} V_{1,t}^*(t,x) + \nabla V_1^*(t,x) \cdot f(t,x,u_1^\sharp,u_2^\sharp) & = \ L_1(t,x,u_1^\sharp,u_2^\sharp)\,, \\ \\ V_{2,t}^*(t,x) + \nabla V_2^*(t,x) \cdot f(t,x,u_1^\sharp,u_2^\sharp) & = \ L_2(t,x,u_1^\sharp,u_2^\sharp)\,, \end{cases} \tag{6.13}$$

where $u_i^\sharp = u_i^\sharp\Big(t,\,x,\,\nabla V_1(t,x),\,\nabla V_2(t,x)\Big)$. Notice that (6.13) has exactly the same form as (5.8). The difference lies in the functions $u_i^\sharp$, which reflect different was of solving the "one shot" infinitesimal game (6.10). We examine here various types of solutions, discussing the well posedness of the corresponding backward Cauchy problem.

6.1 Nash solutions

Let the assumptions (A2) hold. For any given t, x, q_1, q_2, the Nash equilibrium solution of the one-shot game with payoffs (6.10) was considered at (4.5)-(4.6). As shown in Section 5, the Cauchy problem for the system of Hamilton-Jacobi equations (6.13) is usually ill-posed, in any space dimension ≥ 2.

6.2 Stackelberg solutions

For any given t, x, q_1, q_2, the Stackelberg solution of the one-shot game (6.10), with Player 1 as the leader and Player 2 as follower, is obtained as follows.

For each $u_1 \in U_1$, let

$$u_2^\flat(t,x,q_2,u_1) \ \doteq \ \arg\max_{\omega_2 \in U_2} \ \Big\{q_2 \cdot f(t,x,u_1,\omega_2) - L_2(t,x,u_1,\omega_2)\Big\}\,.$$

Then the pair

$$\begin{cases} u_1^\sharp(t,x,q_1,q_2) & \doteq \ \arg\max_{\omega_1 \in U_1} \ \Big\{q_1 \cdot f\Big(t,x,\omega_1,u_2^\flat(t,x,q_2,\omega_1)\Big) - L_2\Big(t,x,\omega_1,u_2^\flat(t,x,q_2,\omega_1)\Big)\Big\}\,, \\ \\ u_2^\sharp(t,x,q_1,q_2) & \doteq \ u_2^\flat\Big(t,x,q_2,u_1^\sharp(t,x,q_1,q_2)\Big), \end{cases} \tag{6.14}$$

provides a Stackelberg solution to the one-shot game (6.10).

We observe that, in the situation considered in (4.9)-(4.10) where the dynamics and the running payoffs are decoupled, the Nash and the Stackelberg solutions coincide. In general, the two definitions yield different functions $(u_1^\sharp, u_2^\sharp)$. However, by the analysis in Chapter 5, the Cauchy problem (5.8)–(5.10) for the value functions V_1, V_2 will still be ill posed, in a generic case.

6.3 Pareto optima

Next, we consider the case where the pair of functions $(u_1^\sharp, u_2^\sharp)$ provides a Pareto optimal solution to the one-shot games (6.10). Notice that, for a given (t, x, q_1, q_2), the game (6.10) will usually have infinitely many Pareto optimal solutions. In order to select one such solution, we introduce a function $\theta = \theta(t, x, q_1, q_2) \in\ ]0,1[$ and consider the pair $(u_1^\sharp, u_2^\sharp)$ which maximizes the combined payoff $\theta\Phi_1 + (1-\theta)\Phi_2$. That is

$$
\begin{aligned}
(u_1^\sharp, u_2^\sharp) \ \doteq\ \arg\max_{(\omega_1,\omega_2)\in U_1\times U_2} \Big\{&\theta\Big(q_1\cdot f(t,x,\omega_1,\omega_2) - L_1(t,x,\omega_1,\omega_2)\Big) \\
&+(1-\theta)\Big(q_2\cdot f(t,x,\omega_1,\omega_2) - L_2(t,x,\omega_1,\omega_2)\Big)\Big\}.
\end{aligned}
\tag{6.15}
$$

Next, assume that the pair of value functions (V_1, V_2) provide a solution to the corresponding system of Hamilton-Jaconi equations (6.13). As in Section 5, we wish to study the hyperbolicity of the linearized system (5.12). Assuming that the argmax in (6.15) is attained in the interior of the admissible set $U_1 \times U_2$, the necessary conditions yield

$$
\theta\left(\nabla V_1\cdot\frac{\partial f}{\partial u_1} - \frac{\partial L_1}{\partial u_1}\right) + (1-\theta)\left(\nabla V_2\cdot\frac{\partial f}{\partial u_1} - \frac{\partial L_2}{\partial u_1}\right) \ =\ 0\,, \tag{6.16}
$$

$$
\theta\left(\nabla V_1\cdot\frac{\partial f}{\partial u_2} - \frac{\partial L_1}{\partial u_2}\right) + (1-\theta)\left(\nabla V_2\cdot\frac{\partial f}{\partial u_2} - \frac{\partial L_2}{\partial u_2}\right) \ =\ 0\,. \tag{6.17}
$$

For convenience, introduce the quantities

$$
a_{jk} \ \doteq\ \nabla V_j\cdot\frac{\partial f}{\partial u_k} - \frac{\partial L_j}{\partial u_k} \qquad\qquad j,k = 1,2\,.
$$

Observe that the identities (6.16)-(6.17) imply

$$
a_{11} \ =\ \frac{\theta-1}{\theta}\,a_{21}\,, \qquad\qquad a_{12} \ =\ \frac{\theta-1}{\theta}\,a_{22}\,. \tag{6.18}
$$

In matrix notation, this homogeneous linear system (5.12) can be written as

$$
\begin{pmatrix} Z_{1,t} \\ Z_{2,t}\end{pmatrix} + \sum_{\alpha=1}^n A^\alpha \begin{pmatrix} Z_{1,x_\alpha} \\ Z_{2,x_\alpha}\end{pmatrix} = \begin{pmatrix} 0 \\ 0\end{pmatrix}, \tag{6.19}
$$

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

$$
A^\alpha \ =\ \begin{pmatrix}
f_\alpha + a_{11}\dfrac{\partial u_1^\sharp}{\partial q_{1\alpha}} + a_{12}\dfrac{\partial u_2^\sharp}{\partial q_{1\alpha}} & a_{11}\dfrac{\partial u_1^\sharp}{\partial q_{2\alpha}} + a_{12}\dfrac{\partial u_2^\sharp}{\partial q_{2\alpha}} \\[3mm]
a_{21}\dfrac{\partial u_1^\sharp}{\partial q_{1\alpha}} + a_{22}\dfrac{\partial u_2^\sharp}{\partial q_{1\alpha}} & f_\alpha + a_{21}\dfrac{\partial u_1^\sharp}{\partial q_{2\alpha}} + a_{22}\dfrac{\partial u_2^\sharp}{\partial q_{2\alpha}}
\end{pmatrix}. \tag{6.20}
$$

By (6.18) this can be written as

$$
A^\alpha \ =\ \begin{pmatrix} f_\alpha & 0 \\ 0 & f_\alpha\end{pmatrix} + \begin{pmatrix} \dfrac{\theta-1}{\theta} \\ 1\end{pmatrix}\left(a_{21}\frac{\partial u_1^\sharp}{\partial q_{1\alpha}} + a_{22}\frac{\partial u_2^\sharp}{\partial q_{1\alpha}} \qquad a_{21}\frac{\partial u_1^\sharp}{\partial q_{2\alpha}} + a_{22}\frac{\partial u_2^\sharp}{\partial q_{2\alpha}}\right). \tag{6.21}
$$

Therefore, given $\xi = (\xi_1, \ldots, \xi_n)$, we can write

$$A(\xi) \;=\; \sum_{\alpha=1}^{n} \xi_\alpha A^\alpha \;=\; \left(\sum_{\alpha=1}^{n} \xi_\alpha f_\alpha\right) I_2 + \begin{pmatrix} 1-\theta^{-1} \\ 1 \end{pmatrix} \otimes \mathbf{w}(\xi)\,, \tag{6.22}$$

where I_2 denotes the 2×2 identity matrix and

$$\mathbf{w}(\xi) \;\doteq\; \sum_{\alpha=1}^{n} \xi_\alpha \left(a_{21}\frac{\partial u_1^\sharp}{\partial q_{1\alpha}} + a_{22}\frac{\partial u_2^\sharp}{\partial q_{1\alpha}} \qquad a_{21}\frac{\partial u_1^\sharp}{\partial q_{2\alpha}} + a_{22}\frac{\partial u_2^\sharp}{\partial q_{2\alpha}} \right).$$

According to (6.22), every 2×2 matrix $A(\xi)$ can be represented as a multiple of the identity matrix plus a matrix of rank one. Therefore, both of its eigenvalues are real.

As observed in [9], this does not guarantee that the system (6.19) is hyperbolic, because for some choices of ξ the two eigenvalues of $A(\xi)$ may coincide and a basis of eigenvectors may not exist. In any case, solutions to differential games obtained using Pareto optima should not experience the wild instabilities found with Nash solutions, where the eigenvalues of the corresponding matrices $A(\xi)$ can be complex.

6.4 Cooperative-competitive solutions

Finally, we examine the case where at each step the game (6.6) is solved in terms of the co-co solution.

The sum of the two payoffs

$$V_+^{(N)}(\tau_j, x) \;\doteq\; V_1^{(N)}(\tau_j, x) + V_2^{(N)}(\tau_j, x)$$

is now a scalar function which can be determined by backward induction, solving the optimization problems

$$V_+^{(N)}(\tau_{j-1}, x) \;=\; \max_{(\omega_1,\omega_2)\in U_1\times U_2} \left\{ V_+^{(N)}\Big(\tau_j\,,\; x + hf(\tau_j, x, \omega_1, \omega_2)\Big) - h\sum_{i=1,2} L_i(\tau_j, x, \omega_1, \omega_2)\right\}. \tag{6.23}$$

Letting $N\to\infty$, it is well known [24] that the functions $V_+^{(N)}$ converge to the value function $V_+ = V_+(t_0, x_0)$ for the optimal control problem

$$\text{maximize:} \quad \sum_{i=1,2}\left(\psi_i(x(T)) - \int_{t_0}^{T} L_i(t, x(t), u_1(t), u_2(t))\, dt\right), \tag{6.24}$$

subject to

$$x(t_0) = x_0\,, \qquad\qquad \dot x(t) \;=\; f(t, x(t), u_1(t), u_2(t))\,. \tag{6.25}$$

The maximum combined payoff (6.24) is sought over all couples of controls $(u_1, u_2) : [t_0, T] \mapsto U_1 \times U_2$.

On the other hand, the difference between the two payoffs

$$V_-^{(N)}(\tau_j, x) \;\doteq\; V_1^{(N)}(\tau_j, x) - V_2^{(N)}(\tau_j, x)$$

is a scalar function which can also be determined by backward induction. Indeed, for each j, the value $V_-^{(N)}(\tau_{j-1}, x)$ is the value of the zero-sum game with payoff

$$J \;=\; V_-^{(N)}\Big(\tau_j\,,\; x + hf(\tau_j,\, x,\, u_1, u_2)\Big) - h\,\Big[L_1(\tau_j,\, x,\, u_1, u_2) -\, L_2(\tau_j,\, x,\, u_1, u_2)\Big]. \tag{6.26}$$

Player 1 seeks to maximize this payoff, while Player 2 wants to minimize it. We recall that this game always has a value, within the family of mixed strategies. Letting $N \to \infty$, the functions $V_-^{(N)}$ converge to the value function $V_- = V_-(t_0, x_0)$ for the zero-sum differential game with payoff

$$J \;=\; \psi_1(x(T)) - \psi_2(x(T)) - \int_{t_0}^{T} \Big[L_1(t, x(t), u_1(t), u_2(t)) - L_2(t, x(t), u_1(t), u_2(t))\Big]\, dt\,,$$

subject to

$$u_1(t) \in U_1, \quad u_2(t) \in U_2\,,$$
$$\dot x \;=\; f(t, x, u_1, u_2) \qquad\qquad x(t_0) = x_0\,.$$

This value function can be determined by solving a scalar H-J equation

$$W_t + H(t, x, \nabla W) \;=\; 0 \qquad\qquad W(T, x) \;=\; \psi_1(x) - \psi_2(x)\,.$$

Here the hamiltonian function $H(t, x, q)$ is defined as value of the "instantaneous" zero-sum game with payoff

$$J^{(t,x,q)}(\omega_1, \omega_2) \;\doteq\; \Big\{q \cdot f(t, x, \omega_1, \omega_2) - L_1(t, x, \omega_1, \omega_2) + L_2(t, x, \omega_1, \omega_2)\Big\}\,.$$

This game always has a value, within the family of randomized strategies. In the case where a saddle point exists, the following min-max coincides with the max-min and one has the representation

$$H(t, x, q) \;\doteq\; \max_{\omega_1 \in U_1} \min_{\omega_2 \in U_2} J^{(t,x,q)}(\omega_1, \omega_2) \;=\; \min_{\omega_2 \in U_2} \max_{\omega_1 \in U_1} J^{(t,x,q)}(\omega_1, \omega_2)\,.$$

By the previous analysis, the co-co solution to a differential game for two players can be found by solving two scalar Hamilton-Jacobi equations. This solution (in the viscosity sense) is unique, and stable w.r.t. perturbations.

7 Nash equilibrium feedbacks with infinite time horizon

As shown in the previous sections, the search for Nash equilibrium solutions to a differential game on a time interval $[0, T]$ usually leads to an ill posed Cauchy problem, which is mathematically intractable.

As an alternative, in this section we consider games in infinite time horizon: $t \in [0, \infty[$, with exponentially discounted payoffs. Assume that the state of the system evolves according to

$$\dot x \;=\; f(x, u_1, u_2) \qquad\qquad u_1(t) \in U_1 \quad u_2(t) \in U_2\,, \tag{7.1}$$

and that the payoff functions are

$$J_i \;=\; \int_0^{\infty} e^{-\rho t}\, \phi_i\Big(x(t),\, u_1(t),\, u_2(t)\Big)\, dt\,. \tag{7.2}$$

Here the constant $\rho > 0$ is a discount rate. Since the dynamics and the payoff functions do not depend explicitly on time, it is natural to seek a Nash equilibrium solution consisting of time-independent feedbacks.

A pair of functions $x \mapsto u_1^*(x) \in U_1$, $x \mapsto u_2^*(x) \in U_2$ will be called a **Nash equilibrium solution in feedback form** *to the non cooperative game (7.1)-(7.2) provided that:*

(i) The map $u_1^*(\cdot)$ is an optimal feedback, in connection with the optimal control problem for the first player:

$$\text{maximize} \quad \int_0^\infty e^{-\rho t}\,\phi_1\Big(x(t),\, u_1(t),\, u_2^*(x(t))\Big)\,dt \tag{7.3}$$

subject to

$$\dot{x} \;=\; f(x, u_1, u_2^*(x)) \qquad\qquad u_1(t) \in U_1\,. \tag{7.4}$$

(ii) The map $u_2^*(\cdot)$ is an optimal feedback, in connection with the optimal control problem for the second player:

$$\text{maximize} \quad \int_0^\infty e^{-\rho t}\,\phi_2\Big(x(t),\, u_1^*(x(t)),\, u_2\Big)\,dt \tag{7.5}$$

subject to

$$\dot{x} \;=\; f(x, u_1^*(x), u_2) \qquad\qquad u_2(t) \in U_2\,. \tag{7.6}$$

A general procedure to find these equilibrium feedbacks $u_1^*(\cdot)$, $u_2^*(\cdot)$ relies on the computation of the value functions. In analogy with (A2), assume:

(A2′) *For any $x \in I\!\!R^n$ and every pair of vectors $(q_1, q_2) \in I\!\!R^n \times I\!\!R^n$, there exist a unique pair $(u_1^\sharp,\, u_2^\sharp) \in U_1 \times U_2$ such that*

$$u_1^\sharp(x, q_1, q_2) \;=\; \operatorname*{argmax}_{\omega\in U_1} \Big\{ q_1 \cdot f(x, \omega, u_2^\sharp) + \phi_1(x, \omega, u_2^\sharp) \Big\}\,, \tag{7.7}$$

$$u_2^\sharp(x, q_1, q_2) \;=\; \operatorname*{argmax}_{\omega\in U_2} \Big\{ q_2 \cdot f(x,\, u_1^\sharp, \omega) + \phi_2(x, u_1^\sharp,\, \omega) \Big\}\,. \tag{7.8}$$

For any $x_0 \in I\!\!R^n$, call $V_1(x_0)$ the maximum payoff for the optimal control problem (7.3)-(7.4), given the initial state $x(0) = x_0$. Similarly, let $V_2(x_0)$ be the maximum payoff for the optimal control problem (7.5)-(7.6), given the initial state $x(0) = x_0$.

On an open region where the value functions V_1, V_2 are continuously differentiable, according to Theorem A.12 in the Appendix these functions satisfy the system of Hamilton-Jacobi equations:

$$\left\{ \begin{array}{rcl} \rho V_1 &=& H^{(1)}(x, \nabla V_1, \nabla V_2)\,, \\ \rho V_2 &=& H^{(2)}(x, \nabla V_1, \nabla V_2)\,, \end{array}\right. \tag{7.9}$$

where, for $i = 1, 2$,

$$H^{(i)}(x, q_1, q_2) \;\doteq\; q_1 \cdot f\Big(x,\, u_1^\sharp(x, q_1, q_2),\, u_2^\sharp(x, q_1, q_2)\Big) + \phi_i\Big(x,\, u_1^\sharp(x, q_1, q_2),\, u_2^\sharp(x, q_1, q_2)\Big)\,.$$

In turn, given a solution V_1, V_2 to the system (7.9), the corresponding feedback controls are obtained as

$$u_i^*(x) \;=\; u_i^\sharp(x,\, \nabla V_1(x),\, \nabla V_2(x)) \qquad\qquad i = 1, 2\,.$$

In general, the system (7.9) is highly nonlinear and difficult to solve. Notice, however, that in the present case we are not looking at a Cauchy problem (which can be ill posed), but at a time-independent problem. For applications, it would already be meaningful to construct solutions to (7.9) on some domain Ω, provided that this domain is *positively invariant* for the corresponding dynamics. In other words, calling $t \mapsto x(t, x_0)$ the solution to

$$\dot x = f(x, u_1^*(x), u_2^*(x)) \qquad\qquad x(0) = x_0\,,$$

the forward invariance property means that

$$x(t, x_0) \in \Omega \qquad\quad \text{whenever} \quad x_0 \in \Omega\,, \quad t \geq 0\,. \tag{7.10}$$

7.1 A perturbation approach

A general theory for systems of Hamilton-Jacobi equations of the form (7.9) is not yet available. To make some progress, one can adopt a perturbation approach. Consider a family of games, depending on a small parameter ε. The dynamics is now

$$\dot x \;=\; f(x, u_1, u_2; \varepsilon)\,, \qquad\qquad u_1(t) \in U_1\,, \quad u_2(t) \in U_2\,, \tag{7.11}$$

while the payoff functions are

$$J_i \;=\; \int_0^\infty e^{-\rho t}\, \phi_i\Big(x(t),\, u_1(t),\, u_2(t);\, \varepsilon\Big)\, dt\,. \tag{7.12}$$

Assume that, for $\varepsilon = 0$, the corresponding system (7.9) is "degenerate", in the sense that can be reduced to a scalar equation. Then one can study what happens in the more general case where ε is not zero, but sufficiently small.

More precisely, assume that, for $\varepsilon = 0$, we are given a pair of Nash equilibrium feedbacks solution $u_1^*(\cdot), u_2^*(\cdot)$, defined for x in an open set $\Omega \subset I\!\!R^n$. Let $\Omega^* \subset \Omega$ be a compact subset with smooth boundary, which is positively invariant for the corresponding dynamics. Namely, assume that

$$\mathbf{n}(x) \cdot f(x, u_1^*(x), u_2^*(x); 0) \;<\; 0 \qquad\qquad \text{for all} \;\; x \in \partial\Omega^*. \tag{7.13}$$

Here $\mathbf{n}(x)$ denotes the unit outer normal to the boundary $\partial\Omega^*$ at the point x.

For $\varepsilon > 0$ sufficiently small, a natural question is whether there exist a unique solution $(V_1^\varepsilon, V_2^\varepsilon)$ (or infinitely many solutions) of the system (7.9), defined on the same domain Ω^* and such that

$$\|V_1^\varepsilon - V_1\|_{\mathcal{C}^0(\Omega^*)} \;\to\; 0\,, \qquad\qquad \|V_2^\varepsilon - V_2\|_{\mathcal{C}^0(\Omega^*)} \;\to\; 0 \qquad\quad \text{as} \;\; \varepsilon \to 0. \tag{7.14}$$

We review here a number of cases where the system (7.9) can be reduced to a scalar equation.

1 - Uncoupled games. Assume that both the dynamics and the payoff functions are completely uncoupled. In other words, assume that the the state variable can be split as $x = (x', x'')$, with

$$\dot{x}' = f_1(x', u_1)\,, \qquad\qquad \dot{x}'' = f_2(x'', u_2)\,,$$

$$J_1 \;=\; \int_0^\infty e^{-\rho t}\,\phi_1\Big(x'(t),\,u_1(t)\Big)\,dt\,,$$

$$J_2 \;=\; \int_0^\infty e^{-\rho t}\,\phi_2\Big(x''(t),\,u_2(t)\Big)\,dt\,.$$

In this case, each player solves his own optimal control problem, entirely independent from the actions of the other player. The value functions $V_1 = V_1(x')$, $V_2 = V_2(x'')$ are found by solving two independent, scalar H-J equations.

In the context of finite time games, small perturbations of this game, introducing some weak coupling between the players, were studied in [25].

2 - One weak player. Consider a family of games depending on a small parameter ε, with dynamics

$$\dot{x} \;=\; f_1(x, u_1) + \varepsilon f_2(x, u_2) \tag{7.15}$$

and payoff functionals

$$J_i = \int_0^\infty e^{-\rho t}\Big(\phi_{i1}(x, u_1) + \phi_{i2}(x, u_2)\Big)\,dt \qquad\qquad i = 1, 2\,. \tag{7.16}$$

Here ε is regarded as the *strength* of the second player. When $\varepsilon = 0$, this player cannot influence in any way the evolution of the system. His optimal strategy is thus the myopic one:

$$u_2 = u_2^*(x) \;=\; \operatorname*{argmax}_{\omega\in U_2}\;\phi_{22}(x, \omega)\,.$$

In this case, the non-cooperative game reduces to an optimal control problem for the first player:

$$\text{maximize} \qquad \int_0^\infty e^{-\rho t}\,\Big(\phi_{11}(x, u_1) + \phi_{12}(x, u_2^*(x))\Big)\,dt\,, \tag{7.17}$$

for a system with dynamics

$$\dot{x} \;=\; f_1(x, u_1). \tag{7.18}$$

As soon as the optimal feedback control $u_1^*(x)$ for the first player is found, this determines in turn the trajectories of the system, and hence the value function for the second player.

In one space dimension, the existence and uniqueness of Nash feedback solutions, for $\varepsilon > 0$ small, was investigated in [6].

3 - Symmetric games. We say that the game (7.1)-(7.2) is *symmetric* if the two players have an identical role, in the sense that

$$f(x, u_1, u_2) = f(x, u_2, u_1), \qquad\qquad U_1 = U_2 = U\,,$$

$$\phi_1(x, u_1, u_2) = \phi_2(x, u_2, u_1)\,.$$

In this situation, it is natural to seek a symmetric solution, with $u_1^*(x) = u_2^*(x)$, $V_1(x) = V_2(x) = V(x)$. The function V can be found by solving the scalar H-J equation

$$\rho V \;=\; \nabla V \cdot f\Big(x,\, u^\sharp(x,\nabla V), u^\sharp(x,\nabla V)\Big) + \phi_i\Big(x,\, u^\sharp(x,\nabla V), u^\sharp(x,\nabla V)\Big)$$

(choosing here $i = 1$ or $i = 2$ does not make a difference). In analogy with (A2), we assume here the existence of a function $u^\sharp$ such that

$$\begin{aligned} u^\sharp(x,q) \;&\doteq\; \operatorname*{argmax}_{\omega\in U} \Big\{ q\cdot f(x,\omega,u^\sharp(x,q)) + \phi_1\Big(x,\omega,u^\sharp(x,q)\Big)\Big\} \\ &=\; \operatorname*{argmax}_{\omega\in U} \Big\{ q\cdot f(x,u^\sharp(x,q),\omega) + \phi_2\Big(x,u^\sharp(x,q),\omega\Big)\Big\}. \end{aligned}$$

For example, assume

$$\begin{aligned} \dot x \;&=\; g(x,u_1) + g(x,u_2), \qquad\qquad u_1(t), u_2(t) \in U\,, \\ J_1 \;&=\; \int_0^\infty e^{-\rho t}\Big[\phi(x,u_1) + \psi(x,u_2)\Big]\,dt\,, \\ J_2 \;&=\; \int_0^\infty e^{-\rho t}\Big[\phi(x,u_2) + \psi(x,u_1)\Big]\,dt\,. \end{aligned}$$

Then

$$u^\sharp(x,q) \;=\; \operatorname*{argmax}_{\omega\in U} \Big\{ q\cdot g(x,\omega) + \phi(x,\omega)\Big\}\,.$$

Assuming that, for $\varepsilon = 0$ the game (7.11)-(7.12) is symmetric, an interesting problem is to investigate the existence of Nash equilibrium feedback solutions for $\varepsilon > 0$ small.

4 - Fully cooperative games. Assume that, for $\varepsilon = 0$, the payoffs (7.12) for the two players coincide, i.e.: $\phi_1 = \phi_2 = \phi(x, u_1, u_2)$. In connection with the dynamics (7.11), we can then consider the optimal control problem

$$\text{maximize:}\quad J(u_1,u_2) \;=\; \int_0^\infty e^{-\rho t}\phi(x,u_1,u_2)\,dt\,,$$

for a controller who can simultaneously choose both controls $u_1(\cdot)$ and $u_2(\cdot)$. A feedback solution $(u_1^*(x),\, u_2^*(x))$ to this optimal control problem yields a Nash equilibrium solution to the original differential game, for $\varepsilon = 0$. Solutions for $\varepsilon > 0$ can be studied by a perturbation approach.

5 - Zero sum games. Assume that, for $\varepsilon = 0$, the game (7.11)-(7.12) is zero-sum, i.e.: $\phi_1 = -\phi_2 = \phi(x, u_1, u_2)$. If

$$\max_{u_1\in U_1}\min_{u_2\in U_2} \Big\{q\cdot f(x,u_1,u_2) + \phi(x,u_1,u_2)\Big\} \;=\; \min_{u_2\in U_2}\max_{u_1\in U_1} \Big\{q\cdot f(x,u_1,u_2) + \phi(x,u_1,u_2)\Big\} \tag{7.19}$$

for every $q \in \mathbb{R}^n$, then the value functions $V_1(x) = -V_2(x) = V(x)$ are well defined. They can be determined by solving the scalar Hamilton-Jacobi equation

$$\rho V \;=\; H(x,\nabla V),$$

where $H(x,q)$ is the common value of the quantities in (7.19). Note that, if the identity (7.19) fails, one can still define $H(x,q)$ as the unique value of the zero-sum game with payoff

$$q \cdot f(x, u_1, u_2) + \phi(x, u_1, u_2) \qquad\qquad u_1 \in U_1\,, \quad u_2 \in U_2\,.$$

By Corollary 2 in Section 2, this value is always well defined, within the family of randomized strategies.

Starting with the feedbacks $u_1^*(\cdot)$, $u_2^*(\cdot)$, known in the case $\varepsilon = 0$, one may study solutions for $\varepsilon > 0$ by a perturbation approach.

8 A game with infinitely many players

In a differential game with a small number of players, each player has the power to modify the state of the system. For example, in an oligopolistic market with a small number of producers, each company can affect the market price by changing its own production level.
In the search for Nash solutions in feedback form, this situation determines severe instabilities, often leading to intractable mathematical problems.

On the other hand, when the number of players is very large, the state of the system is determined by the average behavior: no single player has the power to change the overall evolution of the system. This fact greatly simplifies the mathematical model, and the search for Nash solutions to the differential game.

In recent years, a theory of *mean field games* has emerged, motivated by models in economy and finance, with a large number of agents. As an elementary introduction, we discuss below a specific example, leading to a differential game with infinitely many players. For a comprehensive presentation of mean field games, we refer to the original papers [17, 22].

Consider a game for N competing oil producers, who adjust their production levels as functions of time, in order to maximize their total profits. To model this situation, let us introduce the variables

$$\begin{aligned}
p(t) &= [\text{market price of the oil at time } t],\\
u_i(t) &= [\text{rate at which the } i\text{-th producer extracts the oil, at time } t],\\
c_i(u) &= [\text{cost incurred by the } i\text{-th producer, for extracting the oil at rate } u],\\
\overline{R}_i &= [\text{initial amount of oil in the reserves of the } i\text{-th player}].
\end{aligned}$$

The optimization problem for the i-th player is

$$\text{maximize:} \qquad \int_0^\infty e^{-\rho t}\Big[p(t)u_i(t) - c_i(u_i(t))\Big]\,dt\,, \tag{8.1}$$

subject to

$$\int_0^\infty u_i(t)\,dt \;=\; \overline{R}_i\,. \tag{8.2}$$

To fix the ideas, assume that the costs c_i are quadratic functions of the production rates, say

$$c_i(u) \;=\; a_i u + \frac{b_i}{2} u^2 \,. \tag{8.3}$$

Moreover, assume that at each time t the price $p(t)$ is determined by the market, in terms of a demand function. The demand for oil at time t will be modeled as

$$D(t,p) \;=\; We^{\gamma t} p^{-\alpha}$$

Here $We^{\gamma t}$ represents the total size of the economy, growing at a constant rate γ. The exponent $-\alpha < 0$ accounts for the elasticity of the demand, which shrinks as the price increases.
The price function

$$p \;=\; p\left(t,\; \sum_{i=1}^N u_i\right)$$

is now implicitly determined by the equality between supply and demand, i.e.

$$D(t,\, p(t)) \;=\; \sum_{i=1}^N u_i(t)\,.$$

To derive the optimal strategy for the i-th player, a key observation is that, if

$$u_i \;<<\; \sum_{j=1}^N u_j\,,$$

then the contribution of the i-th player to the determination of the market price is very small. If this holds, then the i-th player can regard the price $p(t)$ as a given function of time, determined by the collective behavior of all other players.

The constraint (8.2) can be taken into account by introducing a Lagrange multiplier. For a given constant λ_i (to be determined later), we thus consider the problem

$$\text{maximize:} \qquad \int_0^\infty \left\{ e^{-\rho t}\Big[p(t)u_i(t) - c_i(u_i(t))\Big] - \lambda_i u_i(t)\right\} dt\,. \tag{8.4}$$

The pointwise optimality conditions are

$$\begin{aligned} u_i(t) \;&=\; \operatorname*{argmax}_{\omega\geq 0} \left\{ e^{-\rho t}\Big(p(t)\omega - c_i(\omega)\Big) - \lambda_i\omega\right\} \\ &=\; \operatorname*{argmax}_{\omega\geq 0} \left\{ e^{-\rho t}\Big(p(t)\omega - \alpha_i\omega - \frac{\beta_i}{2}\omega^2\Big) - \lambda_i\omega\right\} \\ &=\; \frac{1}{\beta_i}\Big[p(t) - \alpha_i - e^{\rho t}\lambda_i\Big]_+\,. \end{aligned} \tag{8.5}$$

If

$$\int_0^\infty \frac{1}{\beta_i}\Big[p(t) - \alpha_i\Big]_+ dt \;\leq\; \overline{R}_i\,,$$

we set $\lambda_i = 0$. Otherwise, we determine the constant $\lambda_i \geq 0$ so that

$$\int_0^\infty \frac{1}{\beta_i}\Big[p(t) - \alpha_i - e^{\rho t}\lambda_i\Big]_+ dt \;=\; \overline{R}_i\,. \tag{8.6}$$

Up to here, this has been a classical derivation of the optimal strategy for one of the producers, provided that he knows in advance the oil price $p(t)$ at all future times. We now discuss how this market price can be predicted.

Assume that at the initial time $t = 0$ there is a large number of producers, of different types. For our purposes, the "type" of a producer is described by a vector $y = (\alpha, \beta, \bar{R})$. This determines his production cost $c(u) = \alpha + \beta u^2/2$ and his initial oil reserves, measured by $\bar{R}$. From the previous derivation, it is clear that, for a given price function $t \mapsto p(t)$, the optimal production strategy $u(\cdot)$ in (8.5) depends only on the type of the producer. We can thus write $u = u(t;\, y,\, p(\cdot))$. Similarly, the Lagrange multiplier $\lambda = \lambda(y,\, p(\cdot))$ in (8.6) depends only on the type $y = (\alpha, \beta, \overline{R})$ of the producer.

To describe how many producers of a given type are present, we think of a continuum of producers, with density function $m(y)$. In other words, given any open set $\Omega \subset I\!R^3$, the number of producers of type $y = (\alpha, \beta, \overline{R}) \in \Omega$ is

$$\int_\Omega m(y)\, dy\,.$$

For a given price function $p(\cdot)$, the total oil production at time t is

$$U(t) \;\doteq\; \int_{I\!R^3} u(t;\, y)\, m(y)\, dy\,. \tag{8.7}$$

At each time t, the market price provides a balance between production and demand, namely

$$U(t) \;=\; D(t,\, p(t))\,.$$

More explicitly, this yields the pointwise identities

$$\int_{I\!R^3} \frac{1}{\beta}\Big[p(t) - \alpha - e^{\rho t}\lambda(y)\Big]_+ m(y)\, dy \;=\; W\, e^{\gamma t} p(t)^{-\alpha} \qquad \text{for all} \;\; t \geq 0\,. \tag{8.8}$$

According to (8.6), the Lagrange multiplier $\lambda = \lambda(\alpha, \beta, \overline{R}) \geq$ is determined by the constraint

$$\int_0^\infty \frac{1}{\beta}\Big[p(t) - \alpha - e^{\rho t}\lambda(y)\Big]_+ dt \;=\; \overline{R} \qquad \text{for all} \;\; y = (\alpha, \beta, \overline{R}) \in Supp(m) \subset I\!R^3\,, \tag{8.9}$$

or else $\lambda = 0$ if

$$\int_0^\infty \frac{1}{\beta}\Big[p(t) - \alpha\Big]_+ dt \;\leq\; \overline{R}\,.$$

Of course, the above identities need to be satisfied only at points y where $m(y) > 0$.

In the end, to solve the optimization problem we need to find two functions: $\lambda(y)$ and $p(t)$, satisfying (8.8)-(8.9). A numerical algorithm computing an approximate solution to these equations is as follows. Construct a family of functions depending on an additional parameter $\theta \in [0, \infty[$, such that

$$\begin{cases} \dfrac{\partial}{\partial\theta} p(t,\theta) \;=\; W\, e^{\gamma t} p(t)^{-\alpha} - \displaystyle\int_{I\!R^3} \frac{1}{\beta}\Big[p(t) - \alpha - e^{\rho t}\lambda(y)\Big]_+ m(y)\, dy\,, \\[2ex] \dfrac{\partial}{\partial\theta}\lambda(y,\theta) \;=\; \displaystyle\int_0^\infty \frac{1}{\beta}\Big[p(t) - \alpha - e^{\rho t}\lambda(y)\Big]_+ dt - \overline{R}\,. \end{cases} \tag{8.10}$$

By guessing a suitable initial data

$$p(t,0) \;=\; p_0(t)\,, \qquad\qquad \lambda(y,0) \;=\; \lambda_0(y)\,, \tag{8.11}$$

we expect that as $\theta \to \infty$ the solution to (8.10) will converge to a steady state. In other words,

$$\bar p(t) \;\doteq\; \lim_{\theta\to\infty} p(t,\theta)\,, \qquad\qquad \bar\lambda(y) \;\doteq\; \lim_{\theta\to\infty} \lambda(y,\theta)$$

should yield the solution to the problem.

Remark 7. In the previous model, as time progresses the players do change type, but only because their reserves shrink. Indeed, the reserve of the i-th player at time t is

$$R_i(t) \;=\; \overline{R}_i - \int_0^t u_i(s)\,ds\,.$$

One could consider other ways for the players to change type:

- A player can buy more oil reserves.
- A player can invest in better technology, and reduce his production costs. In this case, the coefficients α, β in his cost function will decrease.

In addition, one may wish to incorporate stochastic effects.

- Since each player does not know precisely the size of his own oil reserve, his assessment of this reserve may randomly change.
- The market demand for oil may be subject to random fluctuations, due to external factors.

All these aspects can be taken into account, in more realistic (but more complex) models.

9 Appendix

For readers' convenience, we collect here various definitions and basic results of mathematical analysis, which were used as background material.

9.1 Convex sets

A set $A \subseteq \mathbb{R}^n$ is **convex** if, for any two points $x, x' \in A$, the segment that joins them is entirely contained in A. Otherwise stated, A is convex if

$$\theta x + (1-\theta)x' \in A \qquad \text{for every} \quad x, x' \in A, \ \theta \in [0,1] .$$

The **convex hull** of a set A is the smallest convex set which contains A. It can be represented as the set of all convex combinations of elements of A, namely

$$co A \ = \ \left\{ \sum_{i=1}^{N} \theta_i x_i \, ; \qquad N \geq 1, \quad x_i \in A, \quad \theta_i \in [0,1], \quad \sum_{i=1}^{N} \theta_i = 1 \right\}. \tag{9.1}$$

In (9.1), for every integer $N \geq 1$ we are choosing N elements of A and taking their convex combinations. However, when $A \subset \mathbb{R}^n$, taking combinations of $n+1$ elements suffices:

Theorem A.1 (Carathéodory). *Let $A \subseteq \mathbb{R}^n$. Then*

$$co A \ = \ \left\{ \sum_{i=1}^{n+1} \theta_i x_i \, ; \ x_i \in A, \quad \theta_i \in [0,1], \quad \sum_{i=1}^{N} \theta_i = 1 \right\}. \tag{9.2}$$

For a proof, see for example the Appendix in [7].

A function $g : \mathbb{R}^n \mapsto \mathbb{R}$ is **convex** if

$$g\Big(\theta x + (1-\theta) y\Big) \ \leq \ \theta \, g(x) + (1-\theta) \, g(y) \qquad \text{for all} \quad x, y \in \mathbb{R}^n \, , \quad \theta \in [0,1] .$$

We say that g is **concave** if $-g$ is convex.

9.2 Multifunctions

In the following, X denotes a metric space with distance $d(\cdot,\cdot)$. The distance between a point $x \in X$ and set $A \subseteq X$ is defined as the smallest distance between x and points in A, i.e.

$$d(x, A) \ \doteq \ \inf_{a \in A} d(x, a).$$

The open ε-neighborhood around the set A is denoted by

$$B(A, \varepsilon) \ \doteq \ \{x \in X \ : \ d(x, A) < \varepsilon\}.$$

The *Hausdorff distance* between two (nonempty) compact sets $A, A' \subset X$ is defined as

$$d_H(A, A') \ \doteq \ \max \Big\{ d(x, A'), \, d(x', A); \quad x \in A, \ x' \in A' \Big\}.$$

Equivalently, $d_H(A, A')$ can be defined as the infimum of all radii $\rho > 0$ such that A is contained in the ρ-neighborhood around A' while A' is contained in the ρ-neighborhood around A (see figure 11).

$$d_H(A, A') = \inf \Big\{ \rho > 0\,; \quad A \subset B(A', \rho) \quad \text{and} \quad A' \subset B(A, \rho) \Big\}\,.$$

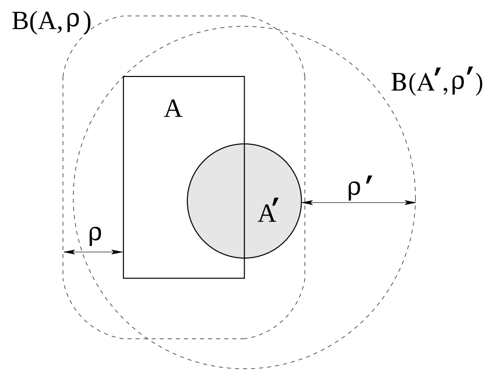

Figure 11: The Hausdorff distance between the two sets A, A' is $\max\{\rho, \rho'\}$.

If X, Y are metric spaces, a *multifunction* F from X to Y is a map that associates to each point $x \in X$ a set $F(x) \subseteq Y$. We say that F is *compact valued* if every $F(x)$ is a non-empty compact subset of Y. The multifunction F is *bounded* if all its values are contained inside a fixed ball $B \subset Y$. We recall here the main continuity concepts for multivalued maps.

Definition. *Let X, Y be metric spaces. A multifunction $F : X \mapsto Y$ with compact values is said to be* **Hausdorff continuous** *if, for every $x \in X$,*

$$\lim_{x' \to x} d_H(F(x'), F(x)) = 0.$$

We say that F is **Hausdorff upper semicontinuous** *if, for every $x \in X$ and $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$F(x') \;\subseteq\; B(F(x), \varepsilon) \qquad\qquad \textit{whenever} \;\; d(x', x) < \delta\,. \tag{9.3}$$

We say that F is **Hausdorff lower semicontinuous** *if, for every $x \in X$ and $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$F(x) \;\subseteq\; B(F(x'), \varepsilon) \qquad\qquad \textit{whenever} \;\; d(x', x) < \delta\,. \tag{9.4}$$

Intuitively, when F is upper semicontinuous, one should think that each set $F(x)$ is "large", compared with the sets $F(x')$ with x' close to x. Indeed, the ε-enlargement $B(F(x), \varepsilon)$ contains all the nearby sets $F(x')$.

On the other hand, when F is lower semicontinuous, the set $F(x)$ is "small" compared with all nearby sets $F(x')$ with x' close to x. Indeed, for all x' in a neighborhood of x, the ε-enlargement $B(F(x'), \varepsilon)$ contains $F(x)$.

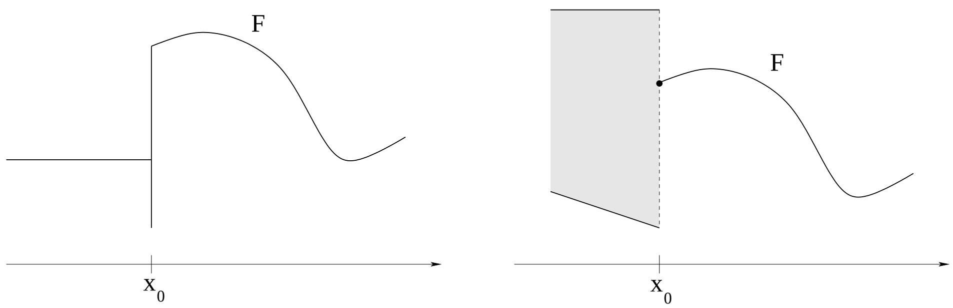

Figure 12: Left: the graph of an upper semicontinuous function. Here $F(x_0)$ contains an entire segment, while $F(x)$ is single-valued for $x \neq x_0$. Right: the graph of a lower semicontinuous function. Here $F(x_0)$ consists of just one point, while $F(x)$ is multivalued for $x < x_0$.

Theorem A.2. *A bounded multifunction $F : X \mapsto \mathbb{R}^N$ with compact values is upper semicontinuous if and only if its graph*

$$Graph(F) = \Big\{(x,y)\,;\quad y \in F(x)\Big\}$$

is closed.

The above condition means that, if $\lim_{k\to\infty} x_k = x$, and $\lim_{k\to\infty} y_k = y$ and $y_k \in F(x_k)$ for every $k \geq 1$, then $y \in F(x)$. For a proof we refer to [2].

Given a multifunction $x \mapsto F(x)$, by a **selection** of F we mean a single-valued function $x \mapsto f(x)$ such that $f(x) \in F(x)$ for every x.

In general, even if F is Hausdorff continuous, in general there may not exists any continuous selection. For convex-valued multifunctions, the main selection theorems are as follows.

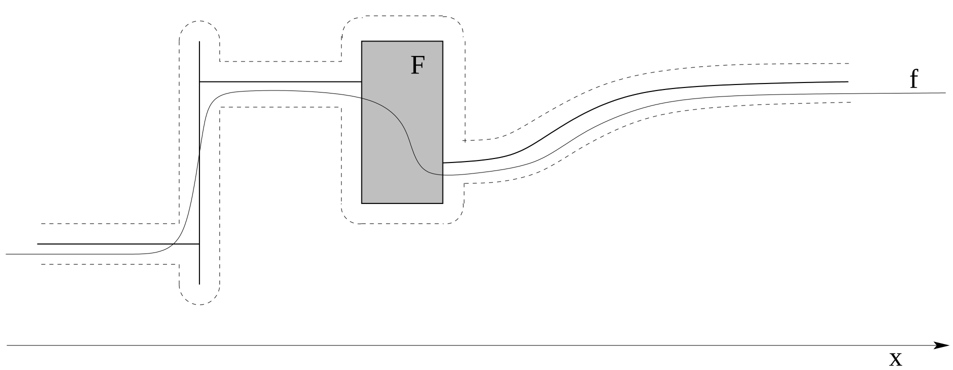

Figure 13: An upper semicontinuous multifunction F with convex values, the ε-neighborhood of its graph, and a continuous ε-approximate selection f.

Theorem A.3 (Michel). *Let X be a metric space, and let $F : X \mapsto \mathbb{R}^N$ be a lower semicontinuous multifunction with compact convex values. Then there exists a continuous selection $x \mapsto f(x) \in F(x)$.*

For a proof we refer to [2].

Theorem A.4 (Cellina). *Let X be a compact metric space, and let $F : X \mapsto I\!R^N$ be an upper semicontinuous multifunction with compact convex values. Then for every $\varepsilon > 0$ there exists a continuous map $f : X \mapsto I\!R^N$ such that*

$$graph(f) \ \subset\ B\Big(graph(F),\, \varepsilon\Big)\,. \tag{9.5}$$

Moreover, f takes values in the convex hull of the image $F(X) \doteq \bigcup_{x\in X} F(x)$.

Proof. Let $\varepsilon > 0$ be given.

1. Since F is upper semicontinuous, for every $x \in X$ there exists a radius $0 < r(x) < \varepsilon/2$ such that

$$F(x') \ \subseteq\ B\Big(F(x),\, \frac{\varepsilon}{2}\Big) \qquad \text{whenever} \qquad |x' - x| < r(x)\,. \tag{9.6}$$

As x ranges in X, the family of all open balls $B(x,\, r(x)/2)$ forms an open covering of the compact metric space X. We can thus choose a finite subcovering, consisting of finitely many balls, say $B(x_\alpha,\, r_\alpha/2)$, $\alpha \in J$.

2. Let $\{\varphi_\alpha\}_{\alpha\in J}$ be a continuous partition of unity subordinated to this covering. For each $\alpha \in J$, we choose a point $y_\alpha \in F(x_\alpha)$. Then we define

$$f(x) \ \doteq\ \sum_{j\in J} \varphi_\alpha(x)\, y_\alpha\,. \tag{9.7}$$

Since this is a finite sum of continuous functions, it is clear that f is a well defined continuous function. Moreover, $f(x) \ \in\ co\{y_\alpha\,;\ \ \alpha \in J\}$. Therefore all values of f lie in the convex hull of the image $F(X)$.

3. To prove (9.5), fix any $x \in X$. Call $J(x) \subseteq J$ the set of all indices such that $\varphi_\alpha(x) > 0$. Choose the largest among all radii r_α, $\alpha \in J(x)$, say

$$r_\beta \ =\ \max_{\alpha\in J(x)} r_\alpha\,.$$

Fore every $\alpha \in J(x)$, this choice implies

$$|x_\alpha - x| \ <\ \frac{r_\alpha}{2} \ \le\ \frac{r_\beta}{2}\,, \qquad\qquad |x_\alpha - x_\beta| \ <\ r_\beta\,.$$

By (9.6) it thus follows

$$y_\alpha \in F(x_\alpha) \ \subseteq\ B\Big(F(x_\beta),\, \frac{\varepsilon}{2}\Big)\,. \tag{9.8}$$

Since f_ε is a convex combination of the points y_α with $\alpha \in J(x)$, and the right hand side of (9.8) is a convex set, this implies

$$f(x) \ \in\ B\Big(F(x_\beta)\,,\, \frac{\varepsilon}{2}\Big).$$

We now compute the distance

$$d\Big((x, f(x)),\ graph(F)\Big) \ \leq\ d(x, x_\beta) + d\Big(f(x),\ F(x_\beta)\Big) \ <\ \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.$$

Since the same inequality holds for every $x \in X$, this proves (9.5). □

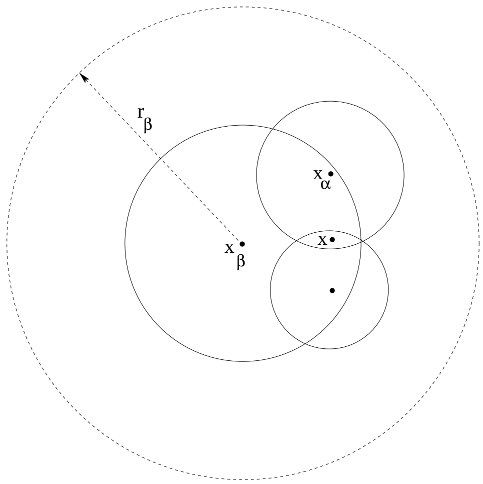

Figure 14: The covering used in Cellina's approximate selection theorem. The ball $B(x_\beta, r_\beta)$ contains all balls $B(x_\alpha, r_\alpha/2)$, for $\alpha \in J(x)$.

9.3 Fixed point theorems

We review here the classical fixed point theorem of Brouwer, and its multivalued extension due to Kakutani.

Theorem A.5 (Brouwer). *Let f be a continuous map from a compact, convex set $K \subset I\!R^n$ into itself. Then there exists a point $\bar{x} \in K$ such that*

$$\bar{x} \ =\ f(\bar{x})\,. \tag{9.9}$$

See for example the Appendix in [7] for an elementary proof.

Corollary A.6 (Kakutani). *Let K be any compact convex subset of $I\!R^n$. Let $F : K \mapsto I\!R^n$ be an upper semicontinuous multifunction with compact, convex values, such that $F(x) \subseteq K$ for every $x \in K$. Then there exists a point $\bar{x} \in K$ such that*

$$\bar{x} \ \in\ F(\bar{x})\,. \tag{9.10}$$

Proof. For every $\varepsilon > 0$, by Cellina's approximate selection theorem there exists a continuous map $f_\varepsilon : K \mapsto K$ such that

$$graph(f_\varepsilon) \subset B(graph(F), \varepsilon).$$

By Brouwer's fixed point theorem, there exists $x_\varepsilon \in K$ such that

$$f(x_\varepsilon) = x_\varepsilon .$$

Since K is compact, we can extract a convergent sequence, say

$$x_{\varepsilon_k} \to \bar{x} .$$

We claim that (9.10) holds. Indeed, by construction

$$d\Big((x_{\varepsilon_k}, f(x_{\varepsilon_k})), \; graph(F)\Big) \leq \varepsilon_k$$

Letting $\varepsilon_k \to 0$ we have

$$(x_{\varepsilon_k}, f(x_{\varepsilon_k})) = (x_{\varepsilon_k}, x_{\varepsilon_k}) \to (\bar{x}, \bar{x}) .$$

Therefore

$$d\Big((\bar{x}, \bar{x}), \; graph(F)\Big) = 0.$$

Hence $(\bar{x}, \bar{x}) \in graph(F)$, because by assumption the graph of F is closed. This implies (9.10). □

9.4 Optimal Control Problems

Here and in the following sections we review some basic definitions and results in the theory of optimal control.

Consider a controlled system described by

$$\dot{x} = f(t, x, u), \qquad u(t) \in U . \tag{9.11}$$

Here t is time, $x \in I\!R^n$ is the state variable, and the upper dot denotes derivative w.r.t. time. The control function $u(\cdot)$ is assumed to be measurable, taking values in a compact domain $U \subset I\!R^m$. Throughout the following we assume

(H1) The function f is continuous w.r.t. all variables and continuously differentiable w.r.t. x. Moreover there exists a constant C such that

$$|f(t, x, u)| \leq C(1 + |x|) \qquad \text{for all } \; (t, x, u) \in [0, T] \times I\!R^n \times U. \tag{9.12}$$

The above linear growth condition guarantees that solutions of the (9.11) cannot become unbounded in finite time. Given an initial condition

$$x(t_0) = x_0 , \tag{9.13}$$

let

$$t \mapsto x(t) = x(t\,; \; t_0, x_0, u) \tag{9.14}$$

be the corresponding trajectory of (9.11), and consider the optimization problem

$$\text{maximize:} \qquad J(u;\, t_0, x_0) \doteq \psi(x(T)) - \int_{t_0}^{T} L(t, x(t), u(t))\, dt . \tag{9.15}$$

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 $u : [t_0, T] \mapsto U$.

9.5 Necessary Conditions for Optimality

Let $t \mapsto u^*(t)$ be an optimal control function, and let $t \mapsto x^*(t) = x(t;\, t_0, x_0, u^*)$ be the corresponding optimal trajectory. A set of necessary conditions satisfied by the functions $u^*(\cdot), x^*(\cdot)$ is provided by the Pontryagin Maximum Principle (PMP). We first consider the basic case where the initial point is given, and the terminal point is free.

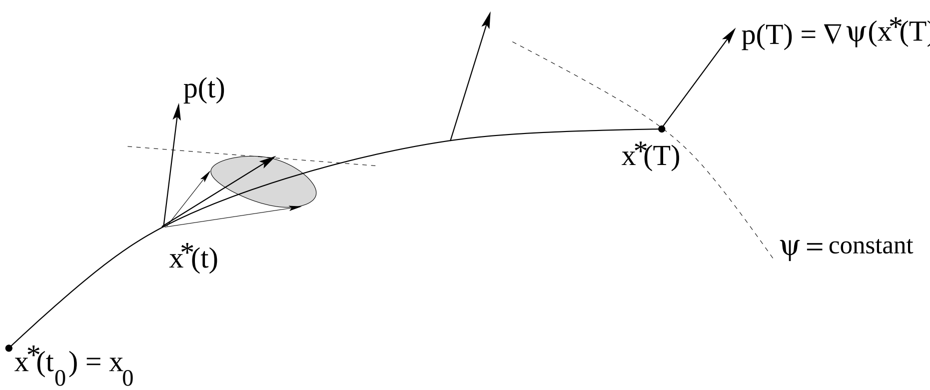

Figure 15: An illustration of the Pontryagin maximum principle, in the case where $L(t, x, u) \equiv 0$. At each time $t \in [0, T]$, among all possible velocities $\dot{x} = f(t, x, u)$, $u \in U$ (shaded set), the optimal choice is the one that maximizes the inner product with the adjoint vector $p(t)$.

Theorem A.7 (PMP, free terminal point). *Let $t \mapsto u^*(t)$ be an optimal control and $t \mapsto x^*(t)$ be the corresponding optimal trajectory for the maximization problem (9.11)–(9.15). Define the vector $t \mapsto p(t)$ as the solution to the linear adjoint system*

$$\dot{p}(t) \;=\; -\,p(t)\,\frac{\partial f}{\partial x}(t, x^*(t),\, u^*(t)) + \frac{\partial L}{\partial x}(t, x^*(t),\, u^*(t)), \tag{9.16}$$

with terminal condition

$$p(T) \;=\; \nabla\psi(x^*(T)). \tag{9.17}$$

Then, for almost every $t \in [t_0, T]$ the following maximality condition holds:

$$p(t)\cdot f(t, x^*(t),\, u^*(t)) - L(t, x^*(t), u^*(t)) \;=\; \max_{u\in U}\Big\{p(t)\cdot f(t, x^*(t),\, u) - L(t, x^*(t), u)\Big\}. \tag{9.18}$$

In the above theorem, x, f, v represent column vectors, while p is a row vector. In coordinates, the above equations (9.16)–(9.18) can be written as

$$\dot{p}_i(t) \;=\; -\sum_{j=1}^{n} p_j(t)\frac{\partial f_j}{\partial x_i}(t, x^*(t), u^*(t)) + \frac{\partial L}{\partial x_i}(t, x^*(t), u^*(t)), \tag{9.19}$$

$$p_i(T) \;=\; \frac{\partial \psi}{\partial x_i}(x^*(T)), \tag{9.20}$$

$$\sum_{i=1}^{n} p_i(t)\cdot f_i(t, x^*(t), u^*(t)) - L(t, x^*(t), u^*(t)) \;=\; \max_{u\in\mathbf{U}}\left\{\sum_{i=1}^{n} p_i(t)\cdot f_i(t, x^*(t), u) \;-\; L(t, x^*(t), u)\right\}. \tag{9.21}$$

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

STEP 1: solve the pointwise maximixation problem (9.18), obtaining the optimal control $u^\sharp$ as a function of t, x, p :

$$u^\sharp(t,x,p) \;=\; \operatorname*{argmax}_{u\in U} \Big\{p\cdot f(t,x,u) - L(t,x,u)\Big\}\,. \tag{9.22}$$

STEP 2: solve the two-point boundary value problem on the interval $[t_0,\,T]$:

$$\begin{cases} \dot x \;=\; f(t,x,\;u^\sharp(t,x,p))\,,\\[2mm] \dot p \;=\; -\,p(t)\,\dfrac{\partial f}{\partial x}(t,x,\;u^\sharp(t,x,p)) + \dfrac{\partial L}{\partial x}(t,x,\;u^\sharp(t,x,p))\,, \end{cases} \qquad \begin{cases} x(t_0) \;=\; x_0\,,\\[2mm] p(T) \;=\nabla\psi(x(T))\,. \end{cases} \tag{9.23}$$

Remark 8. There are two main difficulties associated with the above procedure:

- In general, the function $u^\sharp = u^\sharp(t,p,x)$ in (9.22) is highly nonlinear. It can be multivalued or discontinuous.
- The set of equations (9.23) is not a Cauchy problem, where all components of the solution are explicitly given at the initial time $t = t_0$. On the contrary, the value of x is prescribed at the initial time $t = t_0$, but the value of p is determined by an equation valid at the terminal time $t = T$.

Numerically, the two-point boundary value problem (9.23) can be solved by a *shooting method.* One needs to guess an initial value $\bar p$ and solve the corresponding Cauchy problem, consisting of the system of ODEs in (9.23) with initial data

$$x(t_0) = x_0\,, \qquad\qquad p(t_0) = \bar p\,.$$

This yields a map from $I\!R^n$ into $I\!R^n$:

$$\bar p \;\mapsto\; \Lambda(\bar p) \;\doteq\; p(T) - \nabla\psi(x(T))\,.$$

By an iterative procedure, we then adjust the value of $\bar p$ and try to find a zero of the map Λ. This will provide a solution to the boundary value problem (9.23).

Remark 9. Recalling (9.22), consider the Hamiltonian function

$$H(t,x,p) \;\doteq\; \sup_{u\in U} \Big\{p\cdot f(t,x,u) - L(t,x,u)\Big\} \;=\; p\cdot f(t,x,\;u^\sharp(t,x,p)) - L(t,x,\;u^\sharp(t,x,p)). \tag{9.24}$$

Assume that the functions f, L are continuously differentiable w.r.t. x, u, and assume that the maximum $u^\sharp$ in (9.22) is always attained at an interior point of U. Then the system of differential equations in (9.23) can be written in Hamiltonian form

$$\begin{cases} \dot x \;=\; \dfrac{\partial H}{\partial p}(t,x,p)\,,\\[2mm] \dot p \;=\; -\,\dfrac{\partial H}{\partial x}(t,x,p)\,. \end{cases} \tag{9.25}$$

Indeed, the maximality condition (9.22) implies

$$p \cdot \frac{\partial f}{\partial u} - \frac{\partial L}{\partial u} \;=\; 0.$$

Next, consider the more general optimization problem

$$\text{maximize:} \qquad J \;\doteq\; \varphi(x(0)) + \psi(x(T)) - \int_0^T L(t, x(t), u(t))\, dt\,, \tag{9.26}$$

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 $u : [0,T] \mapsto U$ and all choices of an initial and terminal data, satisfying the constraints

$$x(0) \in S_0\,, \qquad\qquad x(T) \in S_T\,. \tag{9.27}$$

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

Theorem A.8 (PMP, constrained initial and terminal points). *Let $t \mapsto u^*(t)$ be a bounded, optimal control function and let $t \mapsto x^*(t)$ be the corresponding optimal trajectory for the problem (9.26), with dynamics (9.11) and initial and terminal constraints (9.27). Then the following holds.*

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

$$\dot p_i(t) \;=\; -\sum_{j=1}^n p_j(t)\,\frac{\partial f_j}{\partial x_i}(t, x^*(t),\, u^*(t)) + p_0 \frac{\partial L}{\partial x_i}(t, x^*(t),\, u^*(t)) \qquad i = 1, \ldots, n\,. \tag{9.28}$$

(ii) The initial and terminal values of p satisfy

$$\begin{cases} (p_1, \ldots, p_n)(0) &=\; p_0\, \nabla\varphi(x^*(0)) + \mathbf{n}_0\,, \\ (p_1, \ldots, p_n)(T) &=\; p_0\, \nabla\psi(x^*(T)) + \mathbf{n}_T\,. \end{cases} \tag{9.29}$$

for some vector $\mathbf{n}_0$ orthogonal to manifold S_0 at the initial point $x^*(0)$ and some vector $\mathbf{n}_T$ orthogonal to manifold S_T at the terminal point $x^*(T)$.

(ii) The maximality condition

$$\begin{aligned} &\sum_{i=1}^n p_i(t) \cdot f_i(t, x^*(t), u^*(t)) \;-\; p_0\, L(t, x^*(t), u^*(t)) \\ &\qquad = \max_{u \in \mathbf{U}} \left\{ \sum_{i=1}^n p_i(t) \cdot f_i(t, x^*(t), u) \;-\; p_0\, L(t, x^*(t), u) \right\} \end{aligned} \tag{9.30}$$

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

For a proof, see Chapter 5 in [10].

Remark 10. In a standard situation, the sets S_0, S_T are described in terms of a finite number of smooth scalar functions. For example

$$S_0 \doteq \Big\{x \in \mathbb{R}^n\,; \quad \alpha_1(x) = 0, \quad \dots \quad, \alpha_k(x) = 0\Big\}.$$

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

$$\mathbf{n}_0 \;=\; \sum_{i=1}^{k} \lambda_i \,\nabla\alpha_i(x^*(0))$$

for some numbers $\lambda_1, \dots, \lambda_k \in \mathbb{R}$. The first condition in (9.29) can thus be written as

$$(p_1, \dots, p_n)(0) \;=\; p_0 \nabla\varphi(x^*(0)) + \sum_{i=1}^{k} \lambda_i \,\nabla\alpha_i(x^*(0))$$

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

Remark 11. In the special case where the initial point is fixed and the terminal point is free, we have $S_0 = \{x_0\}$ while $S_T = \mathbb{R}^n$. Therefore, $\mathbf{n}_0 \in \mathbb{R}^n$ can be arbitrary while $\mathbf{n}_T = 0$. The boundary conditions for p become

$$(p_1, \dots, p_n)(0) \quad \text{arbitrary}, \qquad\qquad (p_1, \dots p_n)(T) = p_0 \,\nabla\psi(x^*(T)).$$

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 $p_0 = 1$. The necessary conditions in Theorem A.8 thus reduce to those stated in Theorem A.7.

9.6 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**

$$\mathcal{H}(t, x, u, p) \;\doteq\; p \cdot f(t, x, u) - L(t, x, u) \tag{9.31}$$

and the **reduced Hamiltonian**

$$H(t, x, p) \;\doteq\; \max_{\omega \in U} \Big\{p(t) \cdot f(t, x, \omega) - L(t, x, \omega)\Big\}\,. \tag{9.32}$$

Theorem A.9 (PMP + concavity $\Longrightarrow$ optimality). *In the setting of Theorem A.7, consider a measurable function $t \mapsto u^*(t) \in U$ and two absolutely continuous functions $x^*(\cdot), p(\cdot)$*

satisfying the boundary value problem

$$
\begin{cases} \dot x &= f(t,x,\ u^*(t))\,, \\ \dot p &= -\,p(t)\,\dfrac{\partial f}{\partial x}(t,x,\ u^*(t)) + \dfrac{\partial L}{\partial x}(t,x,\ u^*(t))\,, \end{cases} \qquad \begin{cases} x(t_0) &= x_0\,, \\ p(T) &= \nabla\psi(x(T))\,. \end{cases} \tag{9.33}
$$

together with the maximality condition (9.18). Assume that the set U is convex and that the functions

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

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

Proof. Let $u:[t_0,T]\mapsto U$ be any measurable control function. Then

$$
\begin{aligned}
J(u)-J(u^*) &= \psi(x(T))-\psi(x^*(T)) - \int_{t_0}^T \Big[L(t,x(t),u(t)) - L(t,x^*(t),u^*(t))\Big]\,dt \\
&= \psi(x(T))-\psi(x^*(T)) + \int_{t_0}^T \Big\{\Big[\mathcal{H}(t,x(t),p(t),u(t)) - p(t)\,\dot x(t)\Big] \\
&\qquad -\Big[\mathcal{H}(t,x^*(t),p(t),u^*(t)) - p(t)\,\dot x^*(t)\Big]\Big\}\,dt
\end{aligned} \tag{9.34}
$$

Since u^* satisfies the maximality condition (9.18), for a.e. $t\in[t_0,T]$ one has

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

Using these inequalities in (9.34) we obtain

$$
\begin{aligned}
J(u)-J(u^*) &\le \psi(x(T))-\psi(x^*(T)) \\
&\qquad + \int_{t_0}^T \Big\{\Big[H(t,x(t),p(t)) - p(t)\,\dot x(t)\Big] - \Big[H(t,x^*(t),p(t)) - p(t)\,\dot x^*(t)\Big]\Big\}\,dt
\end{aligned} \tag{9.35}
$$

If the map $x\mapsto H(t,x,p(t))$ 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 (9.35) we finally obtain

$$
\begin{aligned}
J(u)-J(u^*) &\le \psi(x(T))-\psi(x^*(T)) - \int_{t_0}^T \Big\{\dot p(t)[x(t)-x^*(t)] + p(t)[\dot x(t)-\dot x^*(t)]\Big\}\,dt \\
&= \psi(x(T))-\psi(x^*(T)) - \Big\{p(T)[x(T)-x^*(T)] - p(t_0)[x(t_0)-x^*(t_0)]\Big\} \\
&\le 0\,.
\end{aligned}
$$

Indeed, the initial and terminal conditions in (9.33), and the concavity of ψ, imply

$$x(t_0) \;=\; x^*(t_0) \;=\; x_0\,, \qquad\qquad \psi(x(T)) \;\leq\; \psi(x^*(T)) + \nabla\psi(x^*(T))[x(T) - x^*(T)]\,.$$

□

9.7 Dynamic Programming

The concavity assumptions used in Theorem A.9 are very restrictive. An alternative approach to the derivation of sufficient optimality conditions relies on the analysis of the value function.

In the optimal control problem (9.11)–(9.15) we regarded the initial data (t_0, x_0) as fixed. However, one can consider a whole family of optimization problems, with variable initial data. Recalling the notations at (9.14)-(9.15), for each $(t_0, x_0) \in [0, T] \times I\!R^N$, let

$$V(t_0, x_0) \;\doteq\; \sup_{u:[t_0,T]\mapsto U} \; J(u; t_0, x_0) \tag{9.36}$$

be the the maximum payoff that can be achieved starting from the state x_0 at time t_0. The function V in (9.36) is called the **value function** for the optimization problem (9.11), (9.15). By (9.13), when $t = T$ one has

$$V(T, x) \;=\; \psi(x)\,. \tag{9.37}$$

A basic property of this value function is:

Theorem A.10 (Principle of Dynamic Programming). *For any initial data $x_0 \in I\!R$ and $0 \leq t_0 < t_1 < T$, one has*

$$V(t_0, x_0) \;=\; \sup_{u:[t_0,t_1]\mapsto U} \left\{ V\Big(t_1,\; x(t_1;\, t_0, x_0, u)\Big) - \int_{t_0}^{t_1} L\Big(t,\; x(t;\, t_0, x_0, u),\; u(t)\Big)\, dt \right\}\,. \tag{9.38}$$

In other words (see fig. 16), 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 (9.38) 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 $[t_0, T]$.

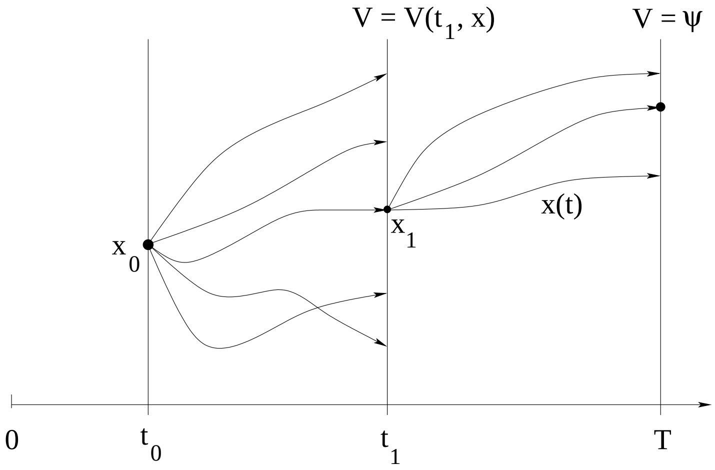

Figure 16: The dynamic programming principle.

Proof on Theorem A.10. Call J_0 the right hand side of (9.38).

1. To prove that $V(t_0,x_0)\leq J_0$, fix $\varepsilon>0$, choose a control $u:[t_0,T]\mapsto U$ such that

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

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

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

Since $\varepsilon>0$ is arbitrary, this first inequality is proved.

2. To prove that $V(t_0,x_0)\geq J_0$, fix $\varepsilon>0$. Then there exists a control $u_0:[t_0,t_1]\mapsto U$ such that

$$V\Big(t_1,\,\,x(t_1,t_0,x_0,u_0)\Big)-\int_{t_0}^{t_1}L(t,\,\,x(t;\,t_0,x_0,u_0),\,\,u_0(t))\,dt\ \geq\ J_0-\varepsilon\,. \tag{9.39}$$

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

$$J(u_1,t_1,x_1)\ \geq\ V(t_1,x_1)-\varepsilon\,. \tag{9.40}$$

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

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

By (9.39) and (9.40) it now follows

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

Since $\varepsilon>0$ can be arbitrarily small, this second inequality is also proved. □

Based on the above property, one can derive a first order PDE satisfied by the value function V, in regions where it is differentiable.

Theorem A.10 (Hamilton-Jacobi-Bellman equation of dynamic programming). *Consider the optimal control problem (9.11)-(9.15), assuming that f, L are continuous functions and that the set U of admissible control values is compact. Then, on any open domain $\Omega \subset [0, T] \times \mathbb{R}^n$ where the function $V = V(t, x)$ is continuously differentiable, the following equation is satisfied:*

$$V_t + H(t, x, \nabla V) \;=\; 0, \tag{9.41}$$

where

$$H(t, x, p) \;\doteq\; \max_{\omega \in U} \Big\{ p \cdot f(t, x, \omega) - L(t, x, \omega) \Big\}. \tag{9.42}$$

Proof. 1. Assume, on the contrary, that at a point $(t_0, x_0) \in \Omega$ one has

$$V_t(t_0, x_0) + \max_{\omega \in U} \Big\{ \nabla V(t_0, x_0) \cdot f(t_0, x_0, \omega) - L(t_0, x_0, \omega) \Big\} \;>\; 0. \tag{9.43}$$

By continuity, we can the find $\omega^* \in U$ and $\theta > 0$ such that

$$V_t(t, x) + \nabla V(t, x) \cdot f(t, x, \omega^*) - L(t, x, \omega^*) \;\geq\; \theta \;>\; 0$$

for all (t, x) in a neighborhood of (t_0, x_0).
Consider the constant control $u^*(t) \equiv \omega^*$, and let $t \mapsto x^*(t) = x(t; t_0, x_0, u^*)$ be the corresponding trajectory. On a sufficiently small interval $[t_0, t_1]$, we have

$$\begin{aligned} V(t_1, x^*(t_1)) \;&=\; V(t_0, x_0) + \int_{t_0}^{t_1} \Big\{ V_t(t, x^*(t)) + \nabla V \cdot f(t, x^*(t), u^*(t)) \Big\}\, dt \\ &\geq\; V(t_0, x_0) + \int_{t_0}^{t_1} \Big\{ \theta + L(t, x^*(t), u^*(t)) \Big\}\, dt \end{aligned}$$

Hence

$$V(t_0, x_0) \;<\; V(t_1, x^*(t_1)) - \int_{t_0}^{t_1} L(t, x^*(t), u^*(t))\, dt$$

against the principle of dynamic programming (9.38). This contradiction shows that (9.43) cannot hold.

2. Next, assume that at a point $(t_0, x_0) \in \Omega$ there holds

$$V_t(t_0, x_0) + \max_{\omega \in U} \Big\{ \nabla V(t_0, x_0) \cdot f(t_0, x_0, \omega) - L(t_0, x_0, \omega) \Big\} \;<\; 0. \tag{9.44}$$

By continuity, it now follows

$$V_t(t, x) + \nabla V(t, x) \cdot f(t, x, \omega) - L(t, x, \omega) \;\leq\; -\theta \;<\; 0$$

for all $\omega \in U$ and all (t, x) in a neighborhood of (t_0, x_0). If the interval $[t_0, t_1]$ is sufficiently small, for any control function $t \mapsto u(t) \in U$ this implies

$$\begin{aligned} V(t_1, x(t_1)) \;&=\; V(t_0, x_0) + \int_{t_0}^{t_1} \Big\{ V_t(t, x(t)) + \nabla V \cdot f(t, x(t), u(t)) \Big\}\, dt \\ &\leq\; V(t_0, x_0) + \int_{t_0}^{t_1} \Big\{ L(t, x(t), u(t)) - \theta \Big\}\, dt. \end{aligned}$$

Therefore, for every admissible control $u(\cdot)$ we have

$$V(t_0, x_0) - \theta(t_1 - t_0) \;\geq\; V(t_1, x(t_1)) - \int_{t_0}^{t_1} L(t, x(t), u(t))\, dt\,,$$

reaching again a contradiction with the principle of dynamic programming (9.38). This completes the proof. □

In cases where the value function V is not $\mathcal{C}^1$, one can still show that the equation (9.41) is satisfied, in a suitable *viscosity sense*. For a comprehensive account of the theory of viscosity solutions we refer to [3].

The PDE of dynamic programming provides sufficient conditions for optimality, in a far more general setting than in Theorem A.9.

Theorem A.11 (sufficient conditions for optimality). *Let $W : [0, T] \times \mathbb{R}^n \mapsto \mathbb{R}$ be a $\mathcal{C}^1$ solution of the terminal value problem*

$$W_t + H(t, x, \nabla W) \;=\; 0, \qquad\qquad W(T, x) \;=\; \psi(x)\,. \tag{9.45}$$

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 (9.11), (9.15).

1. We first show that $V \leq W$.

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

$$J(u, t_0, x_0) \;=\; \psi(x(T)) - \int_{t_0}^{T} L(t, x(t), u(t))\, dt \;\leq\; W(t_0, x_0)\,. \tag{9.46}$$

Observe that

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

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

$$\begin{aligned}
W(t_0, x_0) \;&\geq\; W(T, x(T)) - \textstyle\int_{t_0}^{T} L(t, x(t), u(t))\, dt \\
&=\; \psi(x(T)) - \int_{t_0}^{T} L(t, x(t), u(t))\, dt \;=\; J(u, t_0, x_0)\,.
\end{aligned} \tag{9.47}$$

Therefore,

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

By assumption, the control $u^*(\cdot)$ yields the payoff $J(u^*) = W(t_0,x_0)$. On the other hand, (9.47) 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 in a somewhat easier case, assuming that the map

$$(t,x) \mapsto u^\sharp(t,x) \doteq \operatorname*{argmax}_{u\in U} \Big\{\nabla W(t,x)\cdot f(t,x,u) - L(t,x,u)\Big\}$$

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

$$\dot x = f(t,x,\, u^\sharp(t,x)) \qquad\qquad x(t_0) = x_0\,.$$

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

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

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

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

This establishes the opposite inequality $W \leq V$, completing the proof, at least in the case where the map $u^\sharp$ is continuous. For a proof in more general cases we refer to [11].

9.8 Infinite horizon problems

This section deals with the time-independent control system

$$\dot x = f(x,u) \qquad\qquad u(t)\in U\,, \qquad t\in[0,\infty[\,. \tag{9.48}$$

For a given initial data $x(0) = x_0 \in I\!R^n$ and any admissible control $u : [0,\infty[\mapsto U$, let $t \mapsto x(t) = x(t;x_0,u)$ be the corresponding trajectory. To ensure that this solution is well defined for all times, and does not becomes unbounded in finite time, we assume that f is Lipschitz continuous w.r.t. x and satisfies the bound

$$|f(x,u)| \;\leq\; C(1+|x|) \qquad\qquad \text{for all} \;\; x\in I\!\!R^n, \quad u\in U\,. \tag{9.49}$$

Consider the exponentially discounted payoff

$$J(x_0,u) \;=\; \int_0^\infty e^{-\rho t}\,\Phi(x(t),u(t))\,dt\,, \tag{9.50}$$

where $x(t)=x(t\,;x_0,u)$ and $\rho>0$ is a fixed discount rate. Moreover, define the value function

$$V(x_0) \;\doteq\; \inf_{u(\cdot)} J(x_0,u)\,. \tag{9.51}$$

A result similar to Theorem A.10, valid for infinite horizon problems, is:

Theorem A.12. *Assume that the functions f,Φ are continuous and that $U\subset I\!\!R^m$ is a compact set. Let the value function V be continuously differentiable on an open domain $\Omega\subseteq I\!\!R^n$. Then*

$$\rho V \;=\; H(x,\nabla V(x)) \qquad\qquad \textit{for all} \;\; x\in\Omega\,, \tag{9.52}$$

where the hamiltonian function is

$$H(x,p) \;\doteq\; \max_{\omega\in U}\; \Big\{p\cdot f(x,\omega)+\Phi(x,\omega)\Big\}\,. \tag{9.53}$$

Remark 12. If $W(\cdot)$ is a smooth function satisfying

$$\rho W(x) \;=\; H(x,\,\nabla W(x)) \tag{9.54}$$

for all $x\in I\!\!R^n$, this does not guarantee that W coincides with the value function V at (9.51). To understand the reason, let

$$u \;=\; u^*(x) \;=\; \operatorname*{argmax}_{u\in U}\; \Big\{\nabla W(x)\cdot f(x,\omega)+\Phi(x,\omega)\Big\}$$

be the corresponding optimal feedback control. For a given initial data $x(0)=x_0$, let $t\mapsto x(t)=x(t,x_0)$ be the solution to

$$\dot x = f(x,\,u^*(x))\,, \qquad\qquad x(0)=x_0\,.$$

The identity (9.54) implies

$$\begin{aligned}&\frac{d}{dt}\left[e^{-\rho t}W(x(t))+\int_0^t e^{-\rho s}\Phi\Big(x(s),\,u^*(x(s))\Big)\,ds\right]\\ &\quad=\;-\,\rho W(x(t))+e^{-\rho t}\nabla W(x(t))\cdot f(x(t),u^*(x(t)))+e^{-\rho t}\Phi\Big(x(t),\,u^*(x(t))\Big)\;=\;0\,.\end{aligned}$$

Hence for every $T>0$ we have

$$W(x_0) \;=\; e^{-\rho T}W(x(T))+\int_0^T e^{-\rho t}\Phi\Big(x(t),\,u^*(x(t))\Big)\,dt\,. \tag{9.55}$$

If

$$\lim_{T\to\infty} e^{-\rho T} W(x(T)) \;=\; 0, \tag{9.56}$$

then we let $T \to +\infty$ in (9.55) and conclude

$$W(x_0) \;=\; \int_0^\infty e^{-\rho t}\Phi\Big(x(t),\, u^*(x(t))\Big)\,dt \;\leq\; V(x_0)\,.$$

Notice that limit (9.56) certainly holds if the trajectory $x(\cdot)$ remains bounded. However, one can construct examples where

$$|x(T)| \to \infty, \qquad e^{-\rho T} W(x(T)) \to +\infty\,, \qquad \int_0^T e^{-\rho t}\Phi\Big(x(t),\, u^*(x(t))\Big)\,dt \;\to -\infty$$

as $T \to +\infty$. In this case, one may well have $W(x_0) > V(x_0)$.

9.9 Well posedness for linear PDEs

Consider the Cauchy problem determined by a system of first order PDEs

$$v_t \;=\; G(t, x, v, D_x v) \tag{9.57}$$

together with initial conditions

$$v(0, x) \;=\; \phi(x)\,. \tag{9.58}$$

Here t is time, $x = (x_1, \ldots, x_n)$ is the spatial variable, while $v = (v_1, \ldots, v_m)$ is the dependent variable. By D_x we denote the $m \times n$ Jacobian matrix of partial derivatives $v_{i,x_\alpha} \doteq \partial v_i/\partial x_\alpha$.

A key issue in the theory of PDEs is whether this Cauchy problem is *well posed*. By this we mean that the problem has a unique solution, continuously depending on the initial data ϕ.

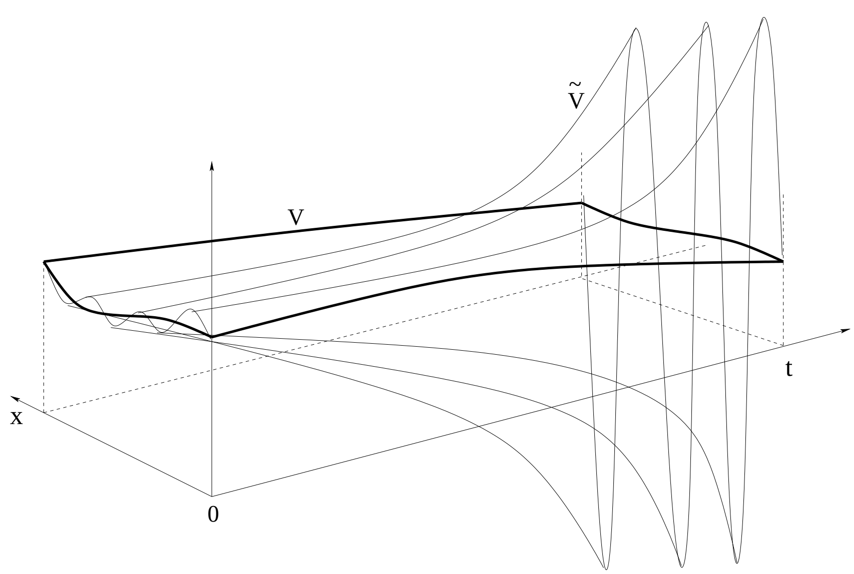

Figure 17: An ill-posed Cauchy problem. Here v is a reference solution, while $\tilde{v}$ is a perturbed solution. A small change in the initial data at $t = 0$ produces a large change in the solution at a later time $t > 0$.

If $m = 1$, then (9.57) reduces to a scalar equation which can be studied by the classical method of characteristics. On the other hand, if $m > 1$, the well-posedness of the system (9.57) is a difficult question, especially if the function G is highly non-linear.

We describe here the method of local linearization, which often provides some useful information.

Let $v = v(t, x)$ be a smooth solution of (9.57), and let

$$v_\varepsilon(t, x) \;=\; v(t, x) + \varepsilon\, Z(t, x) + o(\varepsilon) \qquad (9.59)$$

describe a small perturbation. Assuming that v_ε is also a solution, we can insert (9.59) in the equation (9.57) and compute a system of linearized equations satisfied by Z, namely

$$(Z_j)_t \;=\; \sum_{\alpha=1}^{n}\sum_{k=1}^{m} \frac{\partial G_j}{\partial v_{k,x_\alpha}}\, Z_{k,x_\alpha} \;+\; \sum_{k=1}^{m} \frac{\partial G_j}{\partial v_k}\, Z_k \qquad j = 1, \ldots, m\,. \qquad (9.60)$$

Freezing the coefficients of (9.60) at a given point

$$(\bar t, \bar x, \bar v, \bar Q) \;=\; \Big(\bar t,\ \bar x,\ v(\bar t, \bar x),\ D_x v(\bar t, \bar x)\Big) \;\in\; I\!R \times I\!R^n \times I\!R^m \times I\!R^{m\times n},$$

we obtain a linear system with constant coefficients, namely

$$Z_{j,t} \;=\; \sum_{\alpha=1}^{n}\sum_{k=1}^{m} A^\alpha_{jk}\, Z_{k,x_\alpha} \;+\; \sum_{k=1}^{m} B_{jk} Z_j \qquad j = 1, \ldots, m\,. \qquad (9.61)$$

with

$$A^\alpha_{jk} \;=\; \frac{\partial G_j}{\partial v_{k,x_\alpha}}(\bar t, \bar x, \bar v, \bar Q)\,, \qquad B_{jk} \;=\; \frac{\partial G_j}{\partial v_k}(\bar t, \bar x, \bar v, \bar Q)\,.$$

Notice that each $A^\alpha = (A^\alpha_{jk})$ is a constant $m \times m$ matrix.

Definition. *We say that the Cauchy problem for the linear homogeneous system of PDEs (9.61) is* well posed *in the space* $\mathbf{L}^2(I\!R^n;\, I\!R^m)$ *if the following holds.*

For every initial data $Z(0, \cdot) \in \mathbf{L}^2$*, the equation (9.61) has a unique solution* $t \mapsto Z(t, \cdot)$*. Moreover, for every fixed* $t \in I\!R$*, the map* $Z(0, \cdot) \mapsto Z(t, \cdot)$ *is continuous linear operator from* $\mathbf{L}^2$ *into itself.*

The issue of well-posedness can be studied by means of the Fourier transform. Let $Z = Z(t, x)$ be a solution of (9.61) and let

$$\widehat Z(t, \xi) \;=\; \frac{1}{(2\pi)^{n/2}} \int_{I\!R^n} e^{-i\xi\cdot x} Z(t, x)\, dx \qquad (9.62)$$

be the Fourier transform of Z w.r.t. the spatial variable x. Under suitable regularity and integrability assumptions, taking the Fourier transform of both sides of (9.61) we derive a family of ODEs satisfied by $\widehat Z(\cdot, \xi)$, namely

$$\widehat Z_{j,t}(t, \xi) \;=\; \sum_{\alpha=1}^{n}\sum_{k=1}^{m} i\xi_\alpha A^\alpha_{jk}\, \widehat Z_k(t, \xi) + \sum_{k=1}^{m} B_{jk}\, \widehat Z_k(t, \xi) \qquad j = 1, \ldots, m\,. \qquad (9.63)$$

Adopting a more concise vector notation, this can be written as

$$\widehat{Z}_t(t,\xi) \;=\; (iA(\xi)+B)\widehat{Z}(t,\xi) \tag{9.64}$$

where, for every $\xi = (\xi_1, \ldots, \xi_n)$ we set

$$A(\xi) \;\doteq\; \sum_{\alpha=1}^{n} \xi_\alpha A^\alpha \,.$$

The general solution of (9.64) is given by

$$\widehat{Z}(t,\xi) \;=\; e^{t(iA(\xi)+B)}\,\widehat{Z}(0,\xi)\,. \tag{9.65}$$

As usual, the exponential of an $m \times m$ matrix M is defined as

$$e^M \;\doteq\; \sum_{k=0}^{\infty} \frac{M^k}{k!}\,.$$

Definition. *The system (9.61) is* hyperbolic *if*

$$m_A \;\doteq\; \sup_{\xi\in I\!\!R^n} \|e^{iA(\xi)}\| \;<\; \infty\,. \tag{9.66}$$

Notice that the hyperbolicity of the system depends only on the matrices A^α, not on B. This definition is motivated by

Theorem A.13. *The Cauchy problem for the linear homogeneous system with constant coefficients (9.61) is well posed if and only if the system is hyperbolic.*

For a detailed proof we refer to Chapter 1 in [4]. To understand the underlying idea, assume $B = 0$ and fix any $t \neq 0$. Since the Fourier transform is an isometry on $\mathbf{L}^2$, observing that $tA(\xi) = A(t\xi)$, we compute

$$\begin{aligned} \|Z(t,\cdot)\|_{\mathbf{L}^2} \;&=\; \|\widehat{Z}(t,\cdot)\|_{\mathbf{L}^2} \;\leq\; \|\widehat{Z}(0,\cdot)\|_{\mathbf{L}^2} \cdot \sup_{\xi\in I\!\!R^n} \|e^{iA(t\xi)}\| \\ &=\; \|\widehat{Z}(0,\cdot)\|_{\mathbf{L}^2} \cdot \sup_{\xi\in I\!\!R^n} \|e^{iA(\xi)}\| \;=\; m_A \cdot \|Z(0,\cdot)\|_{\mathbf{L}^2}\,. \end{aligned}$$

2. On the other hand, if the supremum on the left hand side of (9.66) is unbounded, then for every $t > 0$ and any $k > 0$ one can find $\xi^* \in I\!\!R^n$ such that

$$\|e^{iA(t\xi^*)}\| \;>\; k\,.$$

By continuity, there exists $\delta > 0$ such that one still has

$$\|e^{iA(t\xi)}\| \;>\; k \qquad\qquad \text{whenever} \;\; |\xi - \xi^*| < \delta\,.$$

We now consider an initial data $Z(0,x)$ whose Fourier transform satisfies

$$\widehat{Z}(0,\xi) \begin{cases} = 1 & \text{if } |\xi - \xi^*| < \delta\,, \\ = 0 & \text{if } |\xi - \xi^*| \geq \delta\,. \end{cases}$$

This construction yields

$$|\widehat{Z}(t,\xi)| \begin{cases} > k & \text{if } |\xi - \xi^*| < \delta\,, \\ = 0 & \text{if } |\xi - \xi^*| \geq \delta\,. \end{cases}$$

Therefore

$$\|Z(t,\cdot)\|_{\mathbf{L}^2} \;=\; \|\widehat{Z}(t,\cdot)\|_{\mathbf{L}^2} \;\geq\; k\cdot\|\widehat{Z}(0,\cdot)\|_{\mathbf{L}^2} \;=\; k\cdot\|Z(0,\cdot)\|_{\mathbf{L}^2}\,.$$

Since k can be chosen arbitrarily large, the Cauchy problem is not well posed. □

In practice, computing the complex exponential matrix and taking the supremum in (9.66) is a lengthy task. The next theorems provide necessary conditions and sufficient conditions for hyperbolicity, which can be more easily checked.

Theorem A.14 (necessary condition). *If the system (9.61) is hyperbolic, then for every $\xi \in \mathbb{R}^m$ the matrix $A(\xi)$ has a basis of eigenvectors $r_1,\ldots,r_n$, with real eigenvalues $\lambda_1,\ldots,\lambda_n$ (not necessarily distinct).*

Theorem A.15 (sufficient condition). *Assume that, for $|\xi| = 1$, the matrices $A(\xi)$ can be diagonalized in terms of a real, invertible matrix $R(\xi)$ continuously depending on ξ. Then the system (9.61) is hyperbolic.*

For the proofs, we refer to [4].

Example 12. The Cauchy-Riemann equations yield the first order, linear system in one space dimension

$$\begin{cases} u_{1,t} & = & -\,u_{2,x}\,, \\ u_{2,t} & = & u_{1,x}\,. \end{cases}$$

Since the matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has complex eigenvalues, for this system the Cauchy problem is ill posed.

9.10 Probability measures

Let X be a compact metric space and let $\mathcal{C}(X)$ denote the Banach space of all continuous functions $f : X \mapsto \mathbb{R}$, with the standard norm

$$\|f\| \;\doteq\; \max_{x\in X} |f(x)|\,.$$

Let Σ be the σ-algebra of all Borel subsets of X. A **probability measure** on X is a non-negative measure $\mu : \Sigma \mapsto [0,1]$ such that $\mu(X) = 1$.

To each probability measure on X there correspond a unique linear functional on the space $\mathcal{C}(X)$,

$$\Lambda_\mu(f) \doteq \int_X f \, d\mu$$

such that

(i) $|\Lambda_\mu(f)| \leq \|f\|$,

(ii) $f \geq 0 \implies \Lambda_\mu(f) \geq 0$,

(iii) $\Lambda_\mu(\mathbf{1}) = 1$.

Here by "**1**" we denote the constant function identically equal to 1.

We say that a sequence of probability measures $(\mu_k)_{k\geq 1}$ **converges weakly** to a probability measure μ, and write $\mu_k \rightharpoonup \mu$, if for every $f \in \mathcal{C}(X)$ one has

$$\lim_{k\to\infty} \int_X f_k \, d\mu = \int_X f \, d\mu \qquad \text{for all } \ f \in \mathcal{C}(X).$$

Theorem A.16 (weak compactness of the space of probability measures). *Let $(\mu_j)_{j\geq 1}$ be a sequence of probability measures on a compact metric space X. Then there exists a subsequence that converges weakly to some probability measure μ.*

Sketch of the proof. Let $(f_k)_{k\geq 1}$ be a sequence of continuous functions, dense on $\mathcal{C}(X)$. We proceed by induction.

For $k = 1$, by the uniform bound $|\Lambda_{m_k}(f_1)| \leq \|f_1\|$ we can extract a subsequence $(\mu_j)_{j\in J_1}$ such that the sequence $\Lambda_{\mu_j}(f_1)$ converges.

For $k = 2$, by the uniform bound $|\Lambda_{m_k}(f_2)| \leq \|f_2\|$ we can extract a subsequence $(\mu_j)_{j\in J_2}$ with $J_2 \subseteq J_1$ such that the sequence $\Lambda_{\mu_j}(f_1)$ converges.

Continuing the induction procedure, for every k we extract a subsequence $(\mu_j)_{j\in J_k}$ with $J_k \subseteq J_{k-1}$ such that the sequence $\Lambda_{\mu_j}(f_k)$ converges.

Choosing a diagonal sequence $j(1) < j(2) < j(3) < \cdots$ with $j(\ell) \in J_\ell$ for each $\ell \geq 1$, we can define

$$\Lambda_\mu(f) \doteq \lim_{\ell\to\infty} \Lambda_{\mu_{j(\ell}}(f) \qquad f \in \mathcal{C}(X).$$

By construction, the above limit is well defined when f is an element of the dense set $\{f_1, f_2, \ldots\}$. By approximation, one checks that the limit is well defined for every $f \in \mathcal{C}(X)$, and defines a probability measure on X. See [23] for details.

References

[1] J. P. Aubin, *Optima and equilibria.* Second edition. Springer-Verlag, Berlin, 1998.

[2] J. P. Aubin, and A. Cellina, *Differential inclusions. Set-valued maps and viability theory.* Springer-Verlag, Berlin, 1984.

[3] M. Bardi and I. Capuzzo Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Birkhäuser, 1997.

[4] S. Benzoni-Gavage and D. Serre, *Multidimensional hyperbolic partial differential equations. First-order systems and applications.* Oxford University Press, Oxford, 2007.

[5] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory.* Reprint of the second edition. SIAM, Philadelphia, 1999.

[6] A. Bressan, Bifurcation analysis of a noncooperative differential game with one weak player. *J. Differential Equations* **248** (2010), 1297-1314.

[7] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*, AIMS Series in Applied Mathematics, Springfield Mo. 2007.

[8] A. Bressan and W. Shen, Small BV solutions of hyperbolic non-cooperative differential games, *SIAM J. Control Optim.* **43** (2004), 104-215.

[9] A. Bressan and W. Shen, Semi-cooperative strategies for differential games, *Intern. J. Game Theory* **32** (2004), 561-593.

[10] L. Cesari, *Optimization Theory and Applications*, Springer-Verlag, 1983.

[11] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory,* Springer-Verlag, New York, 1998.

[12] A. Cournot, *Recherches sur les principes mathématiques de la théorie des richesses*, L. Hachette, Paris, 1838.

[13] E. J. Dockner, S. Jorgensen, N. V. Long, and G. Sorger, *Differential games in economics and management science.* Cambridge University Press, 2000.

[14] L. C. Evans, *Partial Differential Equations.* Second edition. American Mathematical Society, Providence, RI, 2010.

[15] W.H. Fleming and R.W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York, 1975.

[16] A. Friedman, *Differential Games*, Wiley, New York, 1971.

[17] O. Guéant, J. M. Lasry and P. L. Lions, Mean field games and applications. in: *Paris-Princeton Lectures on Mathematical Finance 2010*, Springer Lecture Notes in Mathematics, Vol. 2003, pp. 205-266.

[18] R. Isaacs, *Differential Games*, Wiley, New York, 1965.

[19] A. T. Kalai and E. Kalai, A cooperative value for Bayesian games, preprint, 2010.

[20] N. N. Krasovskii and A. I. Subbotin, *Game-Theoretical Control Problems* Springer-Verlag, Berlin, 1988.

[21] J. Nash, Non-cooperative games, *Ann. of Math.* **2** (1951), 286-295.

[22] J. M. Lasry and P. L. Lions, Mean field games. *Japanese J. Math.* **2** (2007), 229260.

[23] A. N. Shiryaev, *Probability*, Second edition, Springer-Verlag, New York, 1996.

[24] P. E. Souganidis, Approximation schemes for viscosity solutions of Hamilton-Jacobi equations, *J. Differential Equations* **57** (1985), 1-43.

[25] R. Srikant and T. Basar, Iterative computation of noncooperative equilibria in nonzero-sum differential games with weakly coupled players. *J. Optim. Th. Appl.* **71** (1991), 137-168.

[26] R. B. Vinter, *Optimal Control*, Birkhaüser, Boston, 2000.

[27] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior.* Third edition. Princeton University Press, 1980.

[28] J. Wang, *The Theory of Games.* Oxford University Press, 1988.