# 1 Elements of Optimal Control, Dynamic Programming and Differential Game Theory

The aim of this chapter is to offer a synthetic introduction to optimal control models and differential games, covering the outline of their structure as well as a compact exposition of the solution methods used in applications in the field of industrial organization which appear in the remainder of the book.

Setting out with a summary of a simple dynamic problem featuring a single variable, the exposition is expanded to account for optimal control models with a single agent, and then it is further extended to encompass strategic interaction between at least two players with conflicting objectives, transforming the model into a differential game. Throughout the chapter, the illustration is restricted to noncooperative games; cooperative games are not dealt with, as they are a special case of optimal control models with a single agent controlling several variables.

The concept of Nash equilibrium for differential games is defined. In connection to it, the nature and role of information are discussed under open-loop, closed-loop and feedback rules, to outline the related notions of subgame perfection and weak vs. strong time consistency in differential games. This discussion prompts the analysis of Stackelberg differential games, where the source of time inconsistency, widely known in the macroeconomic policy literature, is identified and a time-consistent solution is outlined.

A summary of the elements of stability analysis based on the properties of the state-control dynamic system and the trace and determinant of the associated Jacobian matrix is also included. More on stability analysis will appear in the context of specific models throughout the book.

For obvious reasons – as is always the case with introductory tutorials about mathematical methods preparatory to the illustration of their applications in any scientific field – this single chapter necessarily falls short of supplying an exhaustive overview of the whole theoretical background, for which the interested reader is referred to the large literature mentioned at the end of the chapter itself.

## 1.1 Preliminaries: The Simplest Dynamic Problem

The point of departure is a dynamic problem with a single object, whose specific nature for the moment is irrelevant, evolving over continuous time. Consider a generic variable $x(t) \in \mathcal{X} \subseteq \mathbb{R}^n$ (which represents a 'state') evolving over continuous time $t$ according to the following ordinary differential equation:

$$\frac{dx}{dt} \equiv \dot{x}(t) = f(x(t), t) \tag{1.1}$$

where $f(x(t), t)$ is continuously differentiable in $x(t)$ and $t$. Such equation is *autonomous* if $\dot{x} = f(x(t))$, i.e., if the time argument does not appear explicitly. If $x(t)$ satisfies (1.1), then it is a solution or an *integral* of the above differential equation. The so-called *Cauchy problem* obtains when we require such a solution to take a specific value $x(0) = x_0$ at the initial instant $t_0$:

$$\begin{cases} \dot{x} = f(x(t), t) \\ x(0) = x_0 \end{cases} \tag{1.2}$$

where $x(0) = x_0$ identifies the *initial condition*. A specific and relevant example of the Cauchy problem describes the evolution of a population, species or, in general, a renewable natural resource as in the Verhulst–Lotka–Volterra model (Verhulst, 1838; Lotka, 1925; Volterra, 1931). Indeed, the presence of the undisturbed resource is in Verhulst (1838). Since we will encounter it in Chapters 2 and 7, the evolution of the model and its basic features can be usefully illustrated here. Suppose the resource is not disturbed by any harvesting activity (i.e., in the traditional jargon of the model, there is no predator), so that its population follows a logistic growth:

$$\dot{x}(t) = zx(t)[1 - vx(t)] \tag{1.3}$$

where $v$ and $z$ are positive parameters. Differential equation (1.3) is in separable variables, and can be easily manipulated to write

$$\frac{dx}{x[1 - vx]} = zdt \Leftrightarrow \frac{dx}{x} + \frac{dx}{\varsigma - x} = zdt \tag{1.4}$$

in which $\varsigma = 1/v$. The next step consists in resorting to logarithms, whereby $\ln(x/x_0) - \ln\left[(\varsigma - x)/(\varsigma - x_0)\right] = zt$. As a result, solving (1.3) yields

$$x^*(t) = \frac{\varsigma e^{zt} x_0}{\varsigma + (e^{zt} - 1) x_0} \tag{1.5}$$

with the asymptotic limit of the population size being $\lim_{t \to \infty} x^*(t) = \varsigma = 1/v$.

If $n$ state variables $\{x_1(t), x_2(t), ...x_n(t)\}$ are present, the Cauchy problem is defined by the dynamic systems

$$\begin{cases} \dot{x}_1 = f_1(x_1(t), x_2(t), ..., x_n(t)) \\ \dot{x}_2 = f_2(x_1(t), x_2(t), ..., x_n(t)) \\ \quad ... \\ \dot{x}_n = f_n(x_1(t), x_2(t), ..., x_n(t)) \end{cases} \tag{1.6}$$

accompanied by a set of $n$ initial conditions $x_i(0) = x_{i0}$, one for each state. Its solution at a generic time $t$ is a vector $\{x_1^*(t), x_2^*(t), ...x_n^*(t)\}$, and a *steady-state equilibrium point* of system (1.6) – which, in general, may not be unique – is identified by coordinates

$$x^{ss} = \left(x_1^{ss}, x_2^{ss}, ...x_n^{ss}\right) \tag{1.7}$$

If there are only two states, we have a planar system whose phase curves can be drawn in the phase plane $(x_1, x_2)$.

Now admit the presence of a predator harvesting the resource, which has thus become a prey. This extension yields the Lotka-Volterra *prey-predator model* (Lotka, 1925; Volterra, 1931). The variable attached to the prey is $x_1$ while that attached to the predator is $x_2$:

$$\begin{cases} \dot{x}_1 = x_1 (\varpi - \beta x_2) \\ \dot{x}_2 = x_2 (\upsilon x_1 - \delta) \end{cases} \tag{1.8}$$

where constants $\{\beta, \delta, \upsilon, \varpi\}$ capture, respectively, (1) the impact of predation (or harvesting activity) on the prey (or resource), (2) the decay rate of the predators; (3) the growth rate of predators given the size of the resource at any time; and (4) the natural growth rate of the resource. The foregoing system has two solutions: $(0, 0)$, involving the extinction of both the preys and predators; and $\left(x_1^* = \delta/\upsilon, x_2^* = \beta/\varpi\right)$, with the two species coexisting in a sustainable way in the long run. The peculiar feature of the prey-predator model is that here neither population literally 'controls' anything, the interaction taking place between two state variables only: it is the pressure exerted by the predators' population onto the preys' one that drives the system towards its steady state, reflecting the idea that here what matters is animal instinct or biology rather than choice. As soon as human beings (and firms) assume the role of predators, conscious and deliberate choices take place and at least a control variable must be added to the model. This is what we will see in Chapters 1 and 7.

## 1.2 Optimal Control Theory

The next step consists in envisaging the realistic possibility that the evolution of states be affected by other variables manoeuvred by one or more agents pursuing explicit objectives. For simplicity, suppose there exist (a) a single state $x(t)$ and (b) a single agent manipulating a single variable; then, (c) define the latter as a *control*, say, $u(t) \in \mathcal{U}$, where $\mathcal{U} \subseteq \mathbb{R}^n$ is the control domain. Additionally, define as (1) $\pi(x(t), u(t), t)$ the instantaneous payoff of the agent controlling $u(t)$, and (2) $f(x(t), u(t), t)$ the function describing the kinematics of state $x(t)$. In the remainder of the chapter, I stipulate that the nature of the model at hand makes it a maximization problem, and that the problem itself is constructed in such a way to meet the concavity conditions.

If the agent controlling $u(t)$ does not discount future payoffs, the control problem defined over a finite time horizon $t \in [0, T]$ consists in

$$\max_{u(t)} \Pi \equiv \int_{t_0}^{T} \pi(x(t), u(t), t)\, dt, \tag{1.9}$$

subject to the state equation

$$\dot{x} = f(x(t), u(t), t) \tag{1.10}$$

and the initial condition $x(t_0) = x_0$, while the terminal condition $x(T)$ is left free, for the moment. Here, $\pi(x(t), u(t), t)$ and $f(x(t), u(t), t)$ are continuously differentiable in $x(t)$, $u(t)$ and $t$. If the relevant time interval is $[t_0, \infty)$, then the optimal control problem is defined over an infinite time horizon.

We can then define the *Hamiltonian function* as

$$\mathcal{H}(x(t), u(t), \mu(t), t) = \pi(x(t), u(t), t) + \mu(t) f(x(t), u(t), t) \tag{1.11}$$

in which $\mu(t)$ is known as the *costate* or *adjoint variable*. The constrained maximisation problem (1.9–1.10) is formally equivalent to maximising Hamiltonian (1.11) s.t. the initial condition $x(t_0) = x_0$. The solution relies on Pontryagin's *Maximum principle* (Pontryagin *et al.*, 1962; Pontryagin, 1966):[1]

**The maximum principle** If $(x^*(t), u^*(t))$ is an optimal couple, then there exists a trajectory $\mu : [t_0, T] \to \mathbb{R}$, not identically equal to zero, such that

- $\dot{\mu} = -\partial\mathcal{H}/\partial x$ where $u^*(t)$ is continuous, and
- the following transversality condition $\mu(T) \geq 0$; $\mu(T) x^*(T) = 0$ is satisfied.

In plain words, solving the Hamiltonian problem requires identifying the optimal instantaneous control $u^*(t)$ and the associated state trajectory. The pair $(x^*(t), u^*(t))$ is called an *optimal couple*. The equation $\dot{\mu} = -\partial\mathcal{H}/\partial x$ is the costate or *adjoint equation*, describing the evolution of the costate. As is the case for the Lagrangian multiplier in static constrained optimization problems, $\mu(t)$ can be thought of as a shadow value (or price). This may be – although not systematically – a sound interpretation of the costate variable that indeed helps intuition in optimal control problems (i.e., with a single agent involved). However, one should refrain from extending this interpretation to differential games, where costates are not, in general, a correct measure of shadow values. More on this very important aspect below.

In general, the value attached to future payoffs is not the same as that of current ones. This is equally true in both economics and politics, and the current debate on environmental values is there to prove the relevance of time discounting.[2] If discounting matters, the constrained optimization problem becomes

$$\max_{u(t)} \Pi \equiv \int_{t_0}^{T} \pi(x(t), u(t), t) e^{-\rho t} dt \tag{1.12}$$

s.t. the same state equation and initial condition as above. In expression (1.12), the payoff flow is discounted at the rate $\rho > 0$. This requires rewriting the Hamiltonian as follows:

[1] Parallel to the work of Pontryagin and associates, Isaacs indentified an equivalent instrument, the *tenet of transition* (Isaacs, 1954, 1965), while working at the RAND Corporation.

[2] See, e.g., Stern (2007, 2009). Although the most common notion of discounting holds that the future counts less than the past – which may be intuitive for firms' profits – there are cases in which it would be wise to think the opposite. And this is not only true for environmental issues, in which the welfare of future generations is at stake. It may also apply when it comes to firms' profits (think of the long-run consequences of advertising campaigns of finite duration) and political parties, who could or should attach a value to the aftermath of electoral campaigns. One such example is presented in Chapter 4.

$$\mathcal{H}(x(t), u(t), \mu(t), t) = \pi(x(t), u(t), t) e^{-\rho t} + \mu(t) f(x(t), u(t), t) \qquad (1.13)$$

or – as it will consistently appear throughout the volume – in its *current value* formulation:

$$\mathcal{H}(x(t), u(t), \mu(t), t) = e^{-\rho t}\left[\pi(x(t), u(t), t) + \lambda(t) f(x(t), u(t), t)\right] \qquad (1.14)$$

in which $\lambda(t) = \mu(t)e^{\rho t}$ is the *capitalised costate variable*. As a result, the adjoint equation becomes

$$\dot{\lambda}(t) = -\frac{\partial \mathcal{H}(\cdot)}{\partial x(t)} + \rho\lambda(t) \qquad (1.15)$$

while the first-order condition (FOC) w.r.t. the control variable is

$$\frac{\partial \mathcal{H}(\cdot)}{\partial u(t)} = e^{-\rho t}\left[\frac{\partial \pi(\cdot)}{\partial u(t)} + \lambda(t) \cdot \frac{\partial f(\cdot)}{\partial u(t)}\right] = 0 \qquad (1.16)$$

The solution derived from conditions (1.15–1.16) must satisfy the transversality condition $\lim_{t\to\infty} x(t)\lambda(t)e^{-\rho t} = 0$.

Before proceeding, a short digression on games played over a finite horizon is in order. If $t \in [0, T]$, the terminal payoff at date $T$ becomes relevant. Call it the *salvage value* and define it as $\mathcal{S}(x(T))$, continuous and differentiable w.r.t. the state. The transversality condition becomes

$$\mu(T) = \frac{\partial \mathcal{S}(\cdot)}{\partial x} \cdot x(T) \qquad (1.17)$$

and becomes $\mu(T) = 0$ if the salvage value is nil or, in the limit, as $T$ tends to infinity. Condition (1.17) tells that the costate cannot be nil at $T$ if there exists a *bequest*, whose interpretation depends on the nature of the specific problem being modelled. For instance, firms may activate investment projects of finite duration, either in advertising, capacity or R&D. Another evident example is that political parties cyclically invest in electoral campaigns whose duration is fixed and known to everybody, including voters. Chapter 4 contains a game of advertising for sale expansion based on this idea.

We can revert to the solution of the optimal control problem. Henceforth I will drop the explicit indication of the time argument, to shorten expressions and simplify the exposition. Manipulating appropriately necessary conditions (1.15–1.16) and the state equation (1.10), one can obtain the control equation

$$\dot{u} = g(x, u) \qquad (1.18)$$

describing the evolution of $u(t)$. Taken together, (1.10) and (1.18) constitute the state-control system

$$\begin{cases} \dot{x} = f(x, u) \\ \dot{u} = g(x, u) \end{cases} \qquad (1.19)$$

describing the dynamics of the optimal control model, and solving (1.19) yields the trajectories of state and control variables in the $(x, u)$ space (plane, if – as here – we have two variables only). If (1.19) is not integrable, one has to linearise the system around the steady state(s) and study the properties of the following $2 \times 2$ *Jacobian matrix*:

$$J = \begin{bmatrix} \dfrac{\partial f(\cdot)}{\partial x} & \dfrac{\partial f(\cdot)}{\partial u} \\ \dfrac{\partial g(\cdot)}{\partial x} & \dfrac{\partial g(\cdot)}{\partial u} \end{bmatrix} \tag{1.20}$$

### 1.2.1 Sketch of the Stability Analysis

The stability properties of the state-control system (1.19) and its stationary points depend on two characteristics (sign and size) of the trace $\mathcal{T}(J)$ and determinant $\Delta(J)$ of the Jacobian matrix (1.20). If

$$\Delta(J) = \frac{\partial f(\cdot)}{\partial x} \cdot \frac{\partial g(\cdot)}{\partial u} - \frac{\partial f(\cdot)}{\partial u} \cdot \frac{\partial g(\cdot)}{\partial x} < 0 \tag{1.21}$$

we have a saddle. That is, the negativity of the determinant is a sufficient condition for the dynamic system to produce a saddle, independently of the sign of the trace of the Jacobian matrix,

$$\mathcal{T}(J) = \frac{\partial f(\cdot)}{\partial x} + \frac{\partial g(\cdot)}{\partial u} \tag{1.22}$$

The assessment of stability (or instability) is a slightly more involved exercise when the determinant is positive. In this case, a stationary point can be

- a stable node, if $\mathcal{T}(J) < 0$ and $\Delta(J) \in \left(0, \mathcal{T}^2(J)/4\right]$;
- an unstable node, if $\mathcal{T}(J) > 0$ and $\Delta(J) \in \left(0, \mathcal{T}^2(J)/4\right]$;
- a stable focus, if $\mathcal{T}(J) < 0$ and $\Delta(J) > \mathcal{T}^2(J)/4$;
- an unstable focus, if $\mathcal{T}(J) > 0$ and $\Delta(J) > \mathcal{T}^2(J)/4$.

The above list includes the most frequent cases (which we will encounter in the remainder of the book) but is not exhaustive. A center is an additional type of steady state, whose nature is such that the characteristic polynomial of the Jacobian matrix admits two complex eigenvalues whose real part is nil. An example of center is one of the steady-state points generated by the Verhulst–Lotka–Volterra model, with coordinates $\left(x_1^* = \delta/\upsilon, x_2^* = \beta/\varpi\right)$.

## 1.3 Dynamic Programming

The alternative technique characterising the *dynamic programming approach* is due to the work of Bellman (1957). This approach permits to solve the entire family of control problems by solving one of them. It does so by introducing the so-called *optimal value function*, assigning the optimal value to each one of many control problems belonging to the same family, and using a solution method based on Bellman's *optimality principle*. To understand the optimality principle, consider the following argument.

Examine a dynamic problem defined over $t \in [0, T]$. Then eliminate part of the time horizon, say, $\left[0, \widetilde{t}\right]$, with $\widetilde{t} < T$. If the solution (or trajectory of the system) solving the initial problem defined for $t \in [0, T]$ is indeed optimal, its portion concerning the residual time interval $\left[\widetilde{t}, T\right]$ must remain optimal when evaluated anew over such residual time horizon, from $\widetilde{t}$ onwards. This means two related and equally important things:

- the optimal solution (trajectory) has to be independent of initial conditions,
- and must be *strongly time consistent*, i.e., robust to a change in the values of initial conditions.

These two requirements, which indeed are satisfied by the dynamic programming approach, amount to saying that the solutions engendered by the optimality principle are subgame (or Markov) perfect. In fact, one can use strong time consistency, Markov perfection or subgame perfection to indicate the same property.

Now we can turn to the solution method. Take $t \in \left[\widetilde{t}, T\right]$ and consider the same objective functional as in the Hamiltonian we have examined earlier, $\Pi = \int_{\widetilde{t}}^{T} \pi\left(x(s), u(s), s\right) e^{-\rho s} ds$. Then define $V(x, t)$ as the value function of the problem at hand. The optimal value function $V^*(x, t)$ must solve the Hamilton–Jacobi–Bellman (HJB) equation

$$-\frac{\partial V(\cdot)}{\partial t} + \rho V(\cdot) = \max_u \left\{ \pi(\cdot) + \frac{\partial V(\cdot)}{\partial x} \cdot f(\cdot) \right\} \tag{1.23}$$

which, more often than not, is simply labelled as the *Bellman equation*. If the time horizon stretches to doomsday $(T \to +\infty)$, one only needs to find $V^*(x)$ and the Bellman equation becomes[3]

$$\rho V(\cdot) = \max_u \left\{ \pi(\cdot) + \frac{\partial V(\cdot)}{\partial x} \cdot f(\cdot) \right\} \tag{1.24}$$

The solution procedure consists in writing the FOC w.r.t. the control variable,

$$\frac{\partial \left[\pi(\cdot) + f(\cdot) \cdot \partial V(\cdot) / \partial x\right]}{\partial u} = 0 \tag{1.25}$$

which (possibly implicitly) identifies the optimal control, $u^*(x)$. If equation (1.25) is explicitly solvable, one can then substitute the expression of the optimal control into either (1.23) or (1.24), so that the Bellman equation is now a function of the state variable only, if the problem is autonomous (otherwise, also, the time argument will pop up explicitly).

In general, we cannot expect to attain a fully analytical solution of either (1.23) or (1.24), unless a reasonable guess about the functional form of the value function can be made. One class of problems in which this is the case is that consisting of linear-quadratic (LQ) models. These are identified by checking that (1) the instantaneous payoff $\pi(\cdot)$ be quadratic in state and control variables, and (2) the evolution of the state (i.e., $f(\cdot)$) be linear in state(s) and control(s). If so, then we may guess that the

[3] This holds only if the problem at hand is time-autonomous (as will consistently be the case throughout the book). This requires the functional forms of both $\pi(\cdot)$ and $f(\cdot)$ to be independent of time.

value function itself is linear-quadratic in the state variable, and can be written as $V(x) = \epsilon_1 x^2 + \epsilon_2 x + \epsilon_3$ – where, if the problem is autonomous, the time argument does not explicitly appear.

If the problem is autonomous and the model is linear-quadratic, the Bellman equation will contain the vector of undetermined parameters $\{\epsilon_1, \epsilon_2, \epsilon_3\}$ appearing in the value function and the state variable $x$, so that the Bellman equation (1.23) can be rewritten in the following form:

$$x^2 \cdot g(\epsilon) + x \cdot h(\epsilon) + \ell(\epsilon) = 0 \tag{1.26}$$

where polynomials $g(\epsilon)$, $h(\epsilon)$ and $\ell(\epsilon)$ contain the vector of undetermined parameters (or just some of them, not necessarily all). Equation (1.26) generates a system of three equations

$$g(\epsilon) = 0; \; h(\epsilon) = 0; \; \ell(\epsilon) = 0 \tag{1.27}$$

which has to be solved w.r.t. $\{\epsilon_1, \epsilon_2, \epsilon_3\}$. Given a model defined in LQ form, the above system yields two solutions consisting of two different triples $\left\{\epsilon_{1i}^*, \epsilon_{2i}^*, \epsilon_{3i}^*\right\}$, $i = I, II$, one stable and the other unstable, although both strongly time consistent. Both of them will be linear in the state $x$, and therefore are usually labelled as *linear feedback solutions*, the meaning of feedback being that they account for the feedback effect of state(s) onto optimal control(s) at any time $t$ of the relevant time span.[4] Each of these feedback solutions can also be differentiated w.r.t. time to yield the control dynamics.

Once the stable solution has been singled out and the corresponding triple $\left\{\epsilon_{1i}^*, \epsilon_{2i}^*, \epsilon_{3i}^*\right\}$ has been substituted back into the value function, this identifies the optimal value function $V^*(x)$, and the Bellman equation is satisfied. Then, the last step consists in solving the differential equation governing the motion of the state variable to fully characterise the steady state point(s).

Before proceeding to the description of differential games, an additional remark is in order concerning the correspondence between the alternative solutions of the same dynamic model with a single agent through optimal control theory and dynamic programming, respectively. To help visualize ideas, suppose the model has a linear-quadratic form. One of the linear feedback solutions generated by the Bellman equation will reproduce the open-loop one generated by the Hamiltonian, but this is not necessarily the stable one. Indeed, we will see both cases in the models reviewed in the next chapter. Now, the problem here is that if the solution of the optimal control problem based on the Hamiltonian function is unstable, this does not imply that the model as such is affected by instability, as there may exist another solution delivered by the dynamic programming approach which is stable but remains out of reach when solving the Hamiltonian formulation of the same problem. I will say more on this and related matter in the remainder of the chapter.

[4] As far as I know, the source of this terminology has a lot to do with electrical and electronical engineering. Feedback is used in audio amplifiers by creating loops feeding back the signal from the power section to the preamplifier section, so as to improve the control of the signal reaching the loudspeakers. As we shall see, loops and feedbacks are extremely relevant as soon as it comes to differential games, where I will carefully discuss the meaning of this terminology in order to avoid risky misunderstandings.

## 1.4 Differential Games with Simultaneous Play

When the model admits the presence of several agents, it becomes a dynamic game – specifically, since we are considering differential equations, a differential game. Let $\mathcal{N} \equiv \{1, 2, 3, ...n\}$ identify the set of players, and suppose $t \in [0, \infty)$.[5] Additionally, $x_i(t) \in \mathcal{X}$ and $u_i(t) \in \mathcal{U}$ define, respectively, player $i$'s state and control variables. The dynamics of $x_i(t)$ is described by the $i$th state equation

$$\dot{x}_i(t) = f_i(\mathbf{x}(t), \mathbf{u}(t)) \tag{1.28}$$

where $\mathbf{x}(t) \equiv \{x_1(t), x_2(t), ...x_n(t)\}$ and $\mathbf{u}(t) \equiv \{u_1(t), u_2(t), ...u_n(t)\}$ are the vectors of state and control variables at any instant $t$. Equation (1.28) says that, in general case, one may expect the dynamics of the $i$th state variable to be affected by all states and controls. Of course, this does not consistently hold true throughout the entire spectrum of games formalising relevant issues in economics, politics and the social sciences in general. Then, one can easily define the vector of initial conditions as $\mathbf{x}(0) \equiv \{x_1(0), x_2(0), ...x_n(0)\} = \mathbf{x}_0$.

Let the instantaneous payoff function of player $i$ be

$$\pi_i(t) = \pi_i(x_i(t), \mathbf{x}_{-i}(t), u_i(t), \mathbf{u}_{-i}(t), t), \tag{1.29}$$

where $\mathbf{x}_{-i}(t)$ and $\mathbf{u}_{-i}(t)$ are, respectively, the vectors of the $n-1$ states and $n-1$ controls pertaining to each player $j \neq i$. Player $i$'s objective is then to maximise the discounted flow of payoffs

$$\Pi_i \equiv \int_0^\infty \pi_i(\cdot)e^{-\rho t}dt \tag{1.30}$$

w.r.t. $u_i(t)$, subject to the set of $n$ constraints (1.28), given the vector of initial conditions $\mathbf{x}_0$ on states, which is assumed to be known to all players alike. All of them use the common and constant discount rate $\rho > 0$.

The optimal strategy defined by each player depends on the information structure characterising the game. To avoid misunderstandings, note that, in general, we will examine games with symmetric and complete information, and 'information structure' has a meaning which is specific to the context and nature of dynamic analysis.

We will consider three different information structures, to which three different equilibrium concepts are associated:

**Definition 1.a (Open-loop information)** *Under open-loop information, the optimal control is $u_i^* = u_i^*(t)$. This means that it is conditional on current time only, and depends on initial conditions $\mathbf{x}_0$.*

In plain words, the adoption of open-loop rules implies that the player decides what to do on the basis of calendar time, regardless of the game's past history and the current values of state variables.

[5] The game can be reformulated in discrete time without significantly affecting its qualitative properties. For further details, see Başar and Olsder (1982, 1995[2]).

**Definition 1.b (Closed-loop memoryless information)** *Under closed-loop memoryless information,* $u_i^* = u_i^*(t, \mathbf{x}(t), \mathbf{x}_0)$ *. This means that the closed-loop memoryless control depends on time, states and initial conditions. Additionally,* $u_i^* = u_i^*(t, \mathbf{x}(t), \mathbf{x}_0)$ *must be continuous in* $t$ *and uniformly Lipschitz in the state vector at any* $t$ *during the game.*

**Definition 1.c (Feedback information)** *Under feedback information,* $u_i^* = u_i^*(t, \mathbf{x}(t))$ *. This means that the feedback strategy depends on time and states at any instant* $t > 0$*, but not initial conditions* $\mathbf{x}_0$*. Additionally, the feedback control must be continuous in* $t$ *and uniformly Lipschitz in* $\mathbf{x}(t)$ *throughout the game.*

Now consider the difference between open-loop, memoryless closed-loop and feedback information. The first implies that a player decides the entire plan of actions at the initial date and then strictly follows it throughout the time horizon of the game. This implies that one must rely on some commitment device (or technology) to fulfil this sort of requirement, without looking at the evolution of the system to see whether the initial plan remains indeed optimal at any intermediate date $\widehat{t} \in [0, \infty)$. I'm asking for a little patience on the part of the reader, before delving into the implications of this aspect of open-loop rules.

Apparently, memoryless closed-loop information solves this problem by requiring players to take into account the states (or stocks) at any point in time. However, the resulting strategy still depends on initial conditions, and therefore any change in the latter implies changes in the resulting closed-loop strategies.

This is where feedback information kicks in, by removing the requirement concerning the role of initial conditions in the design of players' strategies. This might seem a matter of detail, but it is not, as it renders feedback strategies subgame perfect while open-loop and memoryless closed-loop ones are not so, in general (the special cases where this, instead, is true, will be reviewed later here).

I would also like to draw your attention to a terminological misunderstanding affecting here and there the extant literature (in particular, applications of differential game theory), where 'feedback' and 'closed-loop' attributes are somewhat liberally used as synonymous, while in fact they are not if one wants to stick to the appropriate definitions of these two types of information structures. Strictly speaking, feedback strategies are those characterised by solving the HJB equation of a dynamic problem, while closed-loop strategies are all those incorporating loops between states and controls. Hence, a feedback strategy is a closed-loop one, while the opposite is not necessarily true.[6]

Now we can look at the solution concept. Suppose players enjoy complete, symmetric and imperfect information at any point in time, but information becomes complete between any two points in time, i.e., instant after instant. Hence, we are saying that moves are simultaneous at any time $t$, and at that instant all players correctly observe the past history of the game. The resulting definition of Nash equilibrium for a differential game is the following:

[6] Moreover, observe that closed-loop information may take several forms, more or less sophisticated and not equivalent to each other. Here I am only considering the memoryless specification, but there also exist perfect and imperfect closed-loop rules (see Başar and Olsder, 1982, chapter 5).