



# A Crash Course in Differential Games and Applications

Aart de Zeeuw<sup>1,2</sup> 

Accepted: 1 February 2024 / Published online: 25 March 2024  
© The Author(s) 2024

## Abstract

The methodology of differential games is a combination of optimal control theory and game theory. It is the natural framework for economic analysis with strategic interaction and dynamical optimization. The theory gained traction by seminal papers in the early seventies, and it gradually found its way into economics. The purpose of this paper is to make theory and applications of differential games easily accessible by explaining the basics and by developing some characteristic applications. The core of the theory focuses on the open-loop and the multiple Markov-perfect Nash equilibria that use the maximum principle and dynamic programming as the techniques to solve the optimal control problems. The applications are the game of international pollution control and the game of managing a lake, which is an example of an ecological system with tipping points. Finally, it is interesting to note that the discovery of time-inconsistency in the open-loop Stackelberg equilibrium had a huge impact on macroeconomics, since policy making under rational expectations is a Stackelberg differential game.

**Keywords** Differential games · Optimal control · Game theory · International pollution control · Lake system

**JEL Classification** C61 · C72 · Q50

## 1 Introduction

A differential game is the appropriate framework for economic analysis whenever there is dynamical optimization with strategic interaction. Examples of strategic interaction are oligopolistic market structures and international issues. Dynamical optimization or optimal control arises for intertemporal objectives (welfare, utility, or profit) subject to situations that change by the actions of the agents, such as price adjustments, resource extraction and pollution accumulation. Classical examples are duopolistic competition with sticky prices (Fershtman and Kamien 1987), and international pollution control (van der Ploeg

---

✉ Aart de Zeeuw  
A.J.deZeeuw@tilburguniversity.edu

<sup>1</sup> Department of Economics and TSC, Tilburg University, Tilburg, The Netherlands

<sup>2</sup> Beijer Institute of Ecological Economics, Royal Swedish Academy of Sciences, Stockholm, Sweden

and de Zeeuw 1992). Many more examples are presented in a book on differential games and applications (Dockner et al. 2000).

Differential game theory is essentially an extension of optimal control theory to more than one agent. Therefore, the basics is optimal control theory, and game theory enters by providing equilibrium concepts such as the Nash equilibrium and the Stackelberg equilibrium. An important issue is information structure: Do agents have information on the situation and do agents condition their actions on that information? In an optimal control problem without uncertainty, this does not matter but in a game, the equilibrium depends on information structure because strategic interaction changes. Furthermore, in an optimal control problem, the principle of optimality holds (Bellman 1957). This means that the optimal strategy remains optimal if it is recalculated in any state at any point in time. This does not generally hold for the equilibrium strategies of a differential game, but it can be turned into a requirement for equilibrium. This is like the subgame perfectness requirement in extensive-form games (Selten 1975). When the standard techniques for solving optimal control problems are applied to a differential game, this results in an equilibrium without information (i.e., *open loop*), and an equilibrium with information on the current state (i.e., *feedback*) which is perfect. This became the core of differential game analysis. The purpose of this paper is to clarify the methodology of the differential games in a short and transparent way and to present a few interesting applications. Some derivations of results for the game of international pollution control are improved as compared to the original papers. Furthermore, the paper goes beyond Dockner et al. (2000) by including the recent application to the lake game, which is a non-linear differential game with a tipping point.

Game theory usually gives rise to a multiplicity of Nash equilibria. Differential game theory started to note that a multiplicity arises by using different information structures (Başar and Olsder 1982). Driven by two well-known solution techniques for optimal control problems, a focus developed on comparing the open-loop and the feedback or *Markov-perfect* equilibrium. In case of a quadratic objective function and a linear state transition, the open-loop equilibrium is linear, but another type of multiplicity arises for the Markov-perfect equilibrium: besides a linear one, non-linear equilibria exist as well (Tsutsui and Mino 1990). In a game of international pollution control, it is shown that the open-loop equilibrium has the steady state closer to the cooperative outcome than the linear Markov-perfect equilibrium. The intuition is that at the margin each country emits more in case of feedback information because part of an extra unit will be offset by the other countries. However, when the countries employ non-linear strategies, the steady state of the Markov-perfect equilibrium can be close to the cooperative outcome (Dockner and Long 1993). Non-linear strategies can achieve a similar result as the folk theorem for repeated games.

A third multiplicity of Nash equilibria arises when the state transition is non-linear, i.e. when the state transition is sufficiently convex-concave. This type of dynamics became especially relevant with the rise of ecological economics, for management of ecological systems with tipping points (Scheffer et al. 2001). In the prototype model of a lake, the optimal control analysis yields one or two stable steady states. In the case of two steady states, it depends on the initial condition of the lake whether the optimal trajectory ends up in the clean area or in the polluted area of the lake (Mäler et al. 2003). The analysis of symmetric open-loop Nash equilibria yields the same type of dynamics. However, in the case of two stable steady states, multiplicity of open-loop equilibria arises in a range of initial conditions of the lake, with trajectories leading to one or the other steady state. In this range, coordination is the way to settle for the best open-loop equilibrium. If the initial condition is in the polluted area outside this range, then using open-loop strategies means that the game is trapped in the polluted area of the lake.

The analysis of symmetric Markov-perfect Nash equilibria for the lake shows the same type of multiplicity as the non-linear equilibria for a linear-quadratic differential game (Kossioris et al. 2008). An interesting situation occurs if the cooperative outcome has one stable steady state in the clean area of the lake, and the open-loop equilibrium has two stable steady states, one in the clean area and one in the polluted area of the lake. If the initial condition of the lake is such that the lake is trapped in the polluted area by using open-loop strategies, switching to a Markov-perfect equilibrium allows to get out of the polluted area and to end up in a steady state that is close to the steady state of the cooperative outcome. However, the costs remain high.

The theory of differential games studies both Nash and Stackelberg equilibria, but the applications in economics focus mainly on Nash equilibria. However, one result for the Stackelberg equilibrium in differential games had a huge impact on macroeconomics. Stackelberg equilibria have a leader–follower structure or sequential decision structure. It was shown that the open-loop Stackelberg equilibrium is *time-inconsistent*, meaning that the players have an incentive to change their strategies when they can reconsider it after time has passed by (Simaan and Cruz 1973a; b). In macroeconomics, the leader is the government, and the followers are economic agents. The announcement of future government policy breaks down under the rational expectations of the economic agents (Kydland and Prescott 1977). The detection of time-inconsistency in macroeconomic policy was considered sufficiently important to award the Prize in Economic Sciences in memory of Alfred Nobel in 2004.

Section 2 presents some highlights in the history of differential games and applications in economics, with international pollution control as an example. Section 3 discusses differential games in managing an ecological system with tipping points, with the lake game as an example. Section 4 considers Stackelberg equilibria and time-inconsistency in macroeconomics. Section 5 concludes.

## 2 The History of Differential Games

The methodology of differential games is a combination of optimal control theory and game theory. It is the natural framework for dynamical optimization problems where economic agents choose their actions strategically, taking account of the actions of the other agents, and where the situation changes over time due to the actions of all agents. The development of the theory started in the early fifties of last century (Isaacs, 1954–56), with a focus on zero-sum differential games and military applications. The theory gained traction in the late sixties and early seventies of last century by the publication of the seminal papers Starr and Ho (1969a,b) and Simaan and Cruz jr (1973a; b), with a focus on non-zero-sum differential games. This resulted in a textbook that is still widely used (Başar and Olsder 1982). Initially, the economics profession had little interest in differential games, and repeated games dominated the discussion on dynamic games. However, a changing situation, or a changing *state* as it is called in systems and control theory, is such a natural concept in economics that differential games gained attention which resulted in a textbook with many economic applications (Dockner et al. 2000). The *state* is usually a stock like a resource stock, a capital stock, a pollution stock, or a stock of knowledge. Any analysis of strategical interaction with stock-flow dynamics requires a differential games approach.

A repeated game repeats the same game over time, although the history of actions or a stochastic variable may play a role. A differential game explicitly models the transition

of a state, usually in the form of a differential equation. There is a trade-off, however. A repeated game allows for more sophistication in the analysis of strategic interaction, whereas a differential game uses simple game-theoretic equilibrium concepts but has more sophistication in embedding the game in a changing environment. These simple equilibrium concepts are the Nash equilibrium and the Stackelberg equilibrium. If not denoted otherwise, equilibrium means Nash equilibrium in the sequel.

## 2.1 Solution Techniques and Information Structure

Optimal control theory has developed many solution techniques (e.g., see Fleming and Rishel 1975). The best-known techniques for solving an optimal control problem are the *maximum principle* (Pontryagin et al. 1962) and *dynamic programming*, which is based on Bellman's principle of optimality (Bellman 1957). Either the structure of the problem or simply the preference of the researcher determines which technique is used. If both techniques are applicable to the problem at hand, the optimal solution is found either way. However, this is not true for a differential game. The choice of technique in solving the underlying optimal control problems gives rise to different equilibria of the game. One reason is that the maximum principle usually assumes that the strategies are only a function of time whereas dynamic programming assumes, by construction, that the strategies are a function of time and state. By conditioning the strategies on the state, the equilibrium of the game changes. Note that the maximum principle can also assume that the strategies are a function of time and state but in the sequel, this technique is restricted to functions of time only.

The theory of differential games uses the concept of information structure: *open loop* means that the strategies are only a function of time, *closed loop no memory* means that the strategies are a function of time and current state, and *closed loop memory* means that the strategies are a function of time and current as well as past states. An open-loop information structure corresponds to the maximum principle above, and a closed-loop no-memory information structure shows up in dynamic programming. Note, however, that dynamic programming also provides an equilibrium with the perfectness property (introduced by Selten (1975) for extensive-form games): starting at some time in some value of the state, the equilibrium remains the same. It is common in the literature on differential games to refer to a *feedback* equilibrium if dynamic programming is used. In economic applications, it is common to refer to a *Markov-perfect* equilibrium in this case, relating it to perfectness and to the concept of state. Most economic applications compare the open-loop equilibrium, using the maximum principle, with the Markov-perfect equilibrium, using dynamic programming. The reason is that these techniques for solving the underlying optimal control problems are well-known. However, even the closed-loop no-memory information structure already allows for more equilibria, as the next sections will show. The closed-loop memory information structure allows for many more equilibria but in these cases, one cannot rely on standard techniques for solving optimal control problems.

## 2.2 Information Structure and Multiplicity of Equilibria

In the initial periods of the development of differential game theory, a simple example set the stage for the discussion on information structures (Starr and Ho 1969b).

There are two players, two stages, and two actions 0 and 1 for each player at each stage. The game starts in state  $x = 0$  and moves to one of three states  $x = 1$ ,  $x = 2$  and  $x = 3$  in the second stage. The three games in the second stage are the bi-matrix games

$x = 1$	0	1
0	(5,2)	(2,3)
1	<b>(4,1)</b>	(1,4)

$x = 2$	0	1
0	(2,2)	(3,1)
1	(2,4)	<b>(0,3)</b>

$x = 3$	0	1
0	<b>(2,2)</b>	(1,3)
1	(4,1)	(0,2)

where the entries denote *costs*. The Nash equilibria in the second stage are (1,0) with costs (4,1) in state 1, (1,1) with costs (0,3) in state 2, and (0,0) with costs (2,2) in state 3. The game in the first stage is the bi-matrix game

$x = 0$	0	1
0	(2,2)	(2,2)
1	(-1,1)	(5,0)

and the state  $x = 0$  moves to  $x = 1$  by actions (1,0), to  $x = 2$  by actions (0,0) and (1,1), and to  $x = 3$  by actions (0,1).

The two-stage game can be solved in three ways. First, adding the equilibrium results in the second stage to the first stage yields the bi-matrix game

$x = 0$	0	1
0	(2,5)	<b>(4,4)</b>
1	(3,2)	(5,3)

The Nash equilibrium is (0,1) with costs (4,4). The state moves from  $x = 0$  to  $x = 3$ . The costs in the first stage are (2,2), and the costs in the second stage are (2,2). This is the Markov-perfect equilibrium, derived by dynamic programming. The information structure is closed loop no memory. The equilibrium is subgame perfect because it is assumed that Nash equilibria result in each state at the second stage.

Second, if the information structure is open loop, each strategy consists of an action in the first stage and an action in the second stage. Therefore, each player has the strategies (0,0), (0,1), (1,0) and (1,1). This yields the bi-matrix game

$x=0$	00	01	10	11
00	(4,4)	(5,3)	(4,4)	(3,5)
01	(4,6)	(2,5)	(6,3)	(2,4)
10	(4,3)	(1,4)	(7,2)	(8,1)
11	<b>(3,2)</b>	(0,5)	(7,4)	(5,3)

The open-loop Nash equilibrium is (11,00) with total costs (3,2). The state moves from  $x = 0$  to  $x = 1$ . The costs in the first stage are  $(-1,1)$ , and the costs in the second stage are (4,1). The open-loop Nash equilibrium is time-consistent in the sense that (1,0) is the Nash equilibrium in the state  $x = 1$  at the second stage, but it is not subgame perfect because (1,0) is not the Nash equilibrium in the states  $x = 1$  and  $x = 2$ .

The total costs (3,2) in the open-loop equilibrium are lower than the total costs (4,4) in the Markov-perfect equilibrium. The total costs (3,2) are an entry in the bi-matrix game that yields the Markov-perfect equilibrium, for the actions (1,0). In that case, however, the first player has an incentive to deviate and choose action 0, because its total costs 2 are lower than 3. The reason is that the first player knows that the Nash equilibrium is played in the second stage, so that by choosing action 0 and moving to the state  $x = 2$ , its costs in the second stage will be 0. Its costs in the first stage are 2, so that its total costs are 2. In the bi-matrix game for the open-loop equilibrium, the first player does not have this incentive. The reason is that if the second player chooses the strategy (0,0), it does not pay for the first player to choose either (0,0), (0,1) or (1,0), because the total costs 4 are higher than 3. If the first player chooses action 0 in the first stage to move to the state  $x = 2$ , and then action 1, its costs in the second stage will be 2 and not 0, because the second player is committed to choose action 0 in the second stage. In the Markov-perfect equilibrium, the players end up with the total costs (4,4), because the second player also has an incentive to deviate and choose action 1 instead of action 0, because its total costs 4 are lower than 5.

The example shows that the players can be better off if they can commit to an action in the second stage. If the second player commits to action 0 in the second stage, the first player does not have an incentive to deviate from the strategy (1,1). Both players are better off than in case they wait, observe the state, and play the Nash equilibrium in the second stage. However, bygones are bygones. If the players can observe the state, why would they not reconsider their strategies in the second stage, and expect this to happen in the first stage? It is a form of rational expectations. On the other hand, observing the state, processing the information, and reconsidering the strategy may be costly, in which case the open-loop equilibrium remains relevant.

The Markov-perfect equilibrium has a closed-loop no-memory information structure and requires perfectness. It is interesting to consider this information structure without the perfectness requirement. In this case, each strategy consists of an action in the first stage and an action for each possible state in the second stage. Therefore, each player has 16 strategies. However, some of these strategies are not feasible, because actions in the first stage exclude certain states in the second stage. For example, if the first player chooses action 0 in the first stage, the game will not move to  $x = 1$ . It follows that each player has 8 possible strategies. This ultimately yields the third bi-matrix game

$x=0$	000	010	001	011	100	110	101	111
000	(4,4)	(4,4)	(5,3)	(5,3)	(4,4)	(4,4)	(3,5)	(3,5)
010	(4,6)	(4,6)	(2,5)	(2,5)	<b>(4,4)</b>	<b>(4,4)</b>	(3,5)	(3,5)
001	(4,4)	(4,4)	(5,3)	(5,3)	(6,3)	(6,3)	(2,4)	(2,4)
011	(4,6)	(4,6)	(2,5)	(2,5)	(6,3)	(6,3)	(2,4)	(2,4)
100	(4,3)	(1,4)	(4,3)	(1,4)	(7,2)	(8,1)	(7,2)	(8,1)
110	(3,2)	(0,5)	(3,2)	(0,5)	(7,2)	(8,1)	(7,2)	(8,1)
101	(4,3)	(1,4)	(4,3)	(1,4)	(7,4)	(5,3)	(7,4)	(5,3)
111	<b>(3,2)</b>	(0,5)	(3,2)	(0,5)	(7,4)	(5,3)	(7,4)	(5,3)

where the first action denotes the action in the first stage and the second and third action denote the actions for the possible states in the second stage.

The closed-loop no-memory equilibria are (111,000) with total costs (3,2), (010,100) with total costs (4,4), and (010,110) with total costs (4,4). The first one is the open-loop equilibrium, and the third one is the Markov-perfect equilibrium. The second one has the same trajectory as the third one, and thus the same total costs, but prescribes for the second player a different action for the state  $x = 2$ . In this context, equilibria occur that prescribe actions in states that are off the equilibrium path, which moves from  $x = 0$  to  $x = 3$ . It is also clear what commitment is in this context. In the equilibrium (111,000), the first player chooses action 1 and the second player chooses action 0 in the second stage, whatever the state will be. The requirement of perfectness yields the equilibrium (010,110) because in the second stage, (1,1) is the equilibrium for the state  $x = 2$  and (0,0) for the state  $x = 3$ . Note that action 0 of the first player and action 1 of the second player in the first stage both exclude the state  $x = 1$ .

This example makes clear why the focus of differential games turned to open-loop and Markov-perfect equilibria.

### 2.3 Solution Techniques in Detail

Most researchers use the maximum principle (Pontryagin et al. 1962) for solving an optimal control problem. An optimal control problem in continuous time, with control  $u$  and state  $x$ , optimizes an intertemporal objective subject to a differential equation:

$$\max_{u(\cdot)} \left[ \int_0^T e^{-rt} F(x(t), u(t)) dt + e^{-rT} G(x(T)) \right], \quad \text{subject to} \tag{1}$$

$$\dot{x}(t) = f(x(t), u(t)), x(0) = x_0, \tag{2}$$

where  $r$  denotes the discount rate. If the time horizon  $T$  is infinite, the second term in the objective disappears. The maximum principle introduces a co-state  $\lambda$  (a value for the change in the state) and provides necessary conditions in terms of the current-value Hamiltonian function, which is given by  $H(x, u, \lambda) = F(x, u) + \lambda f(x, u)$ , i.e.

$$\begin{aligned} &\max_{u(t)} H(x(t), u(t), \lambda(t)), \\ &\dot{x}(t) = H_\lambda(x(t), u(t), \lambda(t)) = f(x(t), u(t)), x(0) = x_0, \\ &\dot{\lambda}(t) - r\lambda(t) = -H_x(x(t), u(t), \lambda(t)), \lambda(T) = G'(x(T)). \end{aligned} \tag{3}$$

The essence of these necessary conditions is that the dynamical optimization is turned into static optimizations at each time  $t$ , at the expense of a second differential equation (forward looking) in the co-state  $\lambda$ . After solving for the optimal  $u(t)$  and substituting it in the differential equations, a dynamical system in  $(x, \lambda)$  must be solved. This is a two-point boundary value problem, with an initial condition on  $x$  and a condition at  $T$  on  $\lambda$ . If the time horizon  $T$  is infinite, the necessary conditions contain a transversality condition. This usually implies that the optimal path moves towards the steady state of the dynamical system in  $(x, \lambda)$ , along a stable manifold.

The other technique for solving an optimal control problem is dynamic programming, which is based on Bellman’s principle of optimality (Bellman 1957). The core concept of dynamic programming is the *value function*, which is the function of time and state denoting the optimal value starting at that time in that state. The principle of optimality says that reconsidering the optimization, after some time has passed by, does not change the outcome, so that the value function  $W$  satisfies (e.g., see Davis 1977)

$$W(t, x(t)) = \max_{u(\cdot)} \left[ \int_t^{t+\Delta t} e^{-rs} F(x(s), u(s)) ds + W(t + \Delta t, x(t + \Delta t)) \right]. \tag{4}$$

For small  $\Delta t$ , the integral in Eq. (4) can be approximated by  $e^{-rt}F(x(t), u(t))\Delta t$ , and the term  $W(t + \Delta t, x(t + \Delta t))$  by its first-order Taylor expansion, so that  $W(t, x(t))$  cancels out. Dividing by  $\Delta t$  yields a partial differential equation in  $W$ :

$$-W_t(t, x(t)) = \max_{u(t)} [e^{-rt}F(x(t), u(t)) + W_x(t, x(t))f(x(t), u(t))]. \tag{5}$$

Multiplying Eq. (5) with  $e^{rt}$ , and using the current-value function  $V = e^{rt}W$ , with  $V_t = rV + e^{rt}W_t$  and  $V_x = e^{rt}W_x$ , yields

$$rV(t, x(t)) - V_t(t, x(t)) = \max_{u(t)} [F(x(t), u(t)) + V_x(t, x(t))f(x(t), u(t))]. \tag{6}$$

Because the functions  $F$  and  $f$  do not explicitly depend on time  $t$  (but only implicitly via  $x$  and  $u$ ), the problem is stationary. It follows that the function  $V$  depends only on  $x$ , so that  $V_t = 0$ , and Eq. (6) reduces to

$$rV(x) = \max_u [F(x, u) + V'(x)f(x, u)]. \tag{7}$$

It is interesting to see the correspondence between Eq. (7) and the conditions (3) of the maximum principle. By writing  $V' = \lambda$ ,  $u$  maximizes the Hamiltonian function. Substituting the optimal  $u(x)$ , with  $F_u + V'f_u = 0$ , and differentiating with respect to  $x$  yields  $rV' = F_x + V''f + V'f_x$ . Using  $\lambda = V'f$ , it follows that  $\lambda - r\lambda = -(F_x + \lambda f_x)$ . This is not a full mathematical proof, but it is possible to derive a correspondence between the conditions of the maximum principle and dynamic programming. Note, however, that this requires the value function to be twice continuously differentiable.

An example shows the equivalence of these two techniques for solving optimal control problems. Suppose the objective is to maximize the benefits  $\beta y - 0.5y^2$  of production  $y$  corrected for the damage  $0.5\gamma s^2$  of the stock of pollution  $s$ , which results from the accumulation of emissions  $e = y$  that exceed the natural assimilation  $\delta s$ , or



$$\max_{y(\cdot)} \int_0^\infty e^{-rt} [\beta y(t) - 0.5y^2(t) - 0.5\gamma s^2(t)] dt, \quad \text{subject to} \tag{8}$$

$$\dot{s}(t) = y(t) - \delta s(t), s(0) = s_0, \tag{9}$$

where  $r$  denotes the discount rate, and  $\delta$  the natural assimilation rate.

The Hamiltonian  $H(s, y, \lambda) = \beta y - 0.5y^2 - 0.5\gamma s^2 + \lambda(y - \delta s)$  yields the optimality condition  $\beta - y + \lambda = 0$  in (3), and then the dynamical system

$$\begin{aligned} \dot{s}(t) &= \beta - \delta s(t) + \lambda(t), s(0) = s_0, \\ \dot{\lambda}(t) &= \gamma s(t) + (r + \delta)\lambda(t). \end{aligned} \tag{10}$$

The optimal path moves towards the steady state

$$s^* = \frac{\beta(r + \delta)}{\gamma + \delta(r + \delta)}, \lambda^* = \frac{-\beta\gamma}{\gamma + \delta(r + \delta)}. \tag{11}$$

along a stable line in the plane  $(s, \lambda)$ . The variable  $-\lambda$  indicates the marginal cost of emissions, so that the optimal outcome can be implemented by a tax  $\tau = -\lambda$ .

This is a linear-quadratic optimal control problem, with a linear state transition and a quadratic objective function. In such cases, quadratic value functions solve the dynamic programming Eq. (7). Indeed, after some tedious but straightforward calculations, the value function  $V(s) = as^2 + bs + c, a < 0$ , with

$$a = \frac{(r + 2\delta) - \sqrt{(r + 2\delta)^2 + 4\gamma}}{4}, b = \frac{2\beta a}{r + \delta - 2a}, c = \frac{(\beta + b)^2}{2r}, \tag{12}$$

solves Eq. (7). The optimal control is  $y = \beta + 2as + b$ , so that the resulting state transition becomes

$$\dot{s}(t) = (2a - \delta)s(t) + \beta + b, s(0) = s_0. \tag{13}$$

Because  $2a - \delta < 0$ , the differential Eq. (13) is stable, and the state  $s$  converges to the steady state  $s^* = (\beta + b)/(\delta - 2a)$ , which is equal to the steady state in (11) that results from the maximum principle. The two techniques yield the same outcome, but it is clear why most researchers prefer the maximum principle, since the calculations are easier. For a differential game, however, different equilibria result. This is the topic of the next two sections.

Finally, note that production  $y$  and stock of pollution  $s$  are non-negative, so that the analysis should include non-negativity constraints. In this case, it is easy to check this afterwards. However, other constraints on the control and/or the state may play a role and will lead to additional conditions which complicate the analysis.

### 2.4 Open-Loop Equilibrium

In a differential game with  $n$  players, an infinite horizon, and a state  $x$ , each player  $i$  has a control  $u_i$  and optimizes an objective subject to a differential equation:

$$\max_{u_i(\cdot)} \int_0^\infty e^{-rt} F_i(x(t), u_i(t)) dt, i = 1, 2, \dots, n, \quad \text{subject to} \tag{14}$$

$$\dot{x}(t) = f(x(t), u_1(t), \dots, u_n(t)), x(0) = x_0. \tag{15}$$

In the open-loop case, the best response of each player  $i$  to the strategies of the other players is an optimal control problem as in the previous section, with an extra input into the state transition from these other strategies. The maximum principle yields  $n$  sets of necessary conditions as in (3). The open-loop Nash equilibrium requires consistency of the best responses, which implies that these sets of necessary conditions must be solved simultaneously. In case of symmetry, the objectives  $F_i$ , the optimal controls  $u_i$  and the co-states  $\lambda_i$  are the same for  $i = 1, 2, \dots, n$ , which makes the analysis easier.

Extending the example in the previous section to a symmetric differential game (where the stock of pollution accumulates the sum of emissions net of the natural assimilation) yields the Hamiltonians  $H_i(s, y_i, \lambda_i) = \beta y_i - 0.5y_i^2 - 0.5\gamma s^2 + \lambda_i(y_i + \sum_{j \neq i}^n y_j - \delta s)$  with optimality conditions  $\beta - y_i + \lambda_i = 0$ , leading to the dynamical system

$$\begin{aligned} \dot{s}(t) &= n\beta - \delta s(t) + \sum_{i=1}^n \lambda_i(t), s(0) = s_0, \\ \dot{\lambda}_i(t) &= \gamma s(t) + (r + \delta)\lambda_i(t), i = 1, 2, \dots, n. \end{aligned} \tag{16}$$

This is a situation of international pollution control where the players are countries that choose their own production and emission levels  $y_i$  but are only concerned about their own net benefits (van der Ploeg and de Zeeuw 1992). This can be contrasted with the situation in which the countries jointly choose the production and emission levels  $y_i$  to maximize total net benefits. Derivations of the results below are improved as compared to the original publication.

Because the co-states  $\lambda_i$  are the same for  $i = 1, 2, \dots, n$ , (16) is in fact a two-dimensional system, as the system (10). The equilibrium path moves towards the steady state

$$s^{ol} = \frac{n\beta(r + \delta)}{n\gamma + \delta(r + \delta)}, \lambda^{ol} = \frac{-n\beta\gamma}{n\gamma + \delta(r + \delta)}, \tag{17}$$

along a stable line in the plane  $(s, \lambda)$ , where  $ol$  denotes open loop. The variable  $-\lambda^{ol}$  is again the total marginal cost of emissions, so that the countries can implement taxes  $\tau^{ol} = -\lambda^{ol}$  in the steady state of the open-loop equilibrium. If the countries cooperate, the Hamiltonian is  $H(s, y_1, \dots, y_n, \lambda) = \sum_{i=1}^n (\beta y_i - 0.5y_i^2) - 0.5n\gamma s^2 + \lambda(\sum_{i=1}^n y_i - \delta s)$  with the optimality conditions  $\beta - y_i + \lambda = 0$ , leading to the dynamical system

$$\begin{aligned} \dot{s}(t) &= n\beta - \delta s(t) + n\lambda(t), s(0) = s_0, \\ \dot{\lambda}(t) &= n\gamma s(t) + (r + \delta)\lambda(t). \end{aligned} \tag{18}$$

The difference with the system (16) is that the countries now internalize the damage to the other countries ( $n\gamma$ ). The optimal path moves towards the steady state

$$s^{co} = \frac{n\beta(r + \delta)}{n^2\gamma + \delta(r + \delta)}, \lambda^{co} = \frac{-n^2\beta\gamma}{n^2\gamma + \delta(r + \delta)}, \tag{19}$$

along a stable line in the plane  $(s, \lambda)$ , where  $co$  denotes cooperative. It is easy to see that  $s^{co} < s^{ol}$  and  $-\lambda^{co} > -\lambda^{ol}$  for  $n > 1$ . This means that the stock of pollution  $s$  is smaller and the tax  $-\lambda$  on emissions is higher, when the countries cooperate.

### 2.5 Markov-Perfect Equilibria

Using Eq. (7), the dynamic programming equations for the symmetric game of international pollution control in the previous section are

$$rV_i(s) = \max_{y_i} [\beta y_i - 0.5y_i^2 - 0.5\gamma s^2 + V'_i(s)(y_i + \sum_{j \neq i}^n y_j - \delta s)], i = 1, 2, \dots, n. \tag{20}$$

The value functions  $V_i$  are the same for  $i = 1, 2, \dots, n$ . Because this is a linear-quadratic differential game, with a linear state transition and quadratic objective functions, it is to be expected that the quadratic value functions  $V_i(s) = \hat{a}s^2 + \hat{b}s + \hat{c}$ ,  $\hat{a} < 0$ ,  $i = 1, 2, \dots, n$ , solve Eq. (20). Then the optimal controls  $y_i(s) = \beta + 2\hat{a}s + \hat{b}$ ,  $i = 1, 2, \dots, n$ , form the linear Markov-perfect equilibrium. Indeed, after some tedious but straightforward calculations, the value functions  $V_i(s) = \hat{a}s^2 + \hat{b}s + \hat{c}$ , with

$$\begin{aligned} \hat{a} &= \frac{(r + 2\delta) - \sqrt{(r + 2\delta)^2 + 4(2n - 1)\gamma}}{4(2n - 1)}, \\ \hat{b} &= \frac{2n\beta\hat{a}}{r + \delta - 2(2n - 1)\hat{a}}, \hat{c} = \frac{(\beta + \hat{b})(\beta + (2n - 1)\hat{b})}{2r}, \end{aligned} \tag{21}$$

solve Eq. (20). The resulting state transition becomes

$$\dot{s}(t) = (2n\hat{a} - \delta)s(t) + n(\beta + \hat{b}), s(0) = s_0. \tag{22}$$

Because  $2n\hat{a} - \delta < 0$ , the differential Eq. (22) is stable, and the state  $s$  converges to the steady state

$$s^{mp} = \frac{n(\beta + \hat{b})}{\delta - 2n\hat{a}}. \tag{23}$$

The interesting question is how the steady state of the open-loop equilibrium relates to the steady state of the Markov-perfect equilibrium. Note that  $\hat{a}$  is the negative root of  $2(2n - 1)\hat{a}^2 - (r + 2\delta)\hat{a} - 0.5\gamma = 0$ . Using this and (21), Eq. (23) becomes

$$s^{mp} = \frac{n\beta(r + \delta - 2(n - 1)\hat{a})}{(\delta - 2n\hat{a})(r + \delta - 2(2n - 1)\hat{a})} = \frac{n\beta(r + \delta - 2(n - 1)\hat{a})}{n\gamma + \delta(r + \delta) - 2(n - 1)\delta\hat{a}}. \tag{24}$$

Note that comparing (19), (17) and (24) yields  $s^{co} = s^{ol} = s^{mp}$ , for  $n = 1$ , and  $s^{co} < s^{ol}$ , for  $n > 1$ . Furthermore, it is easy to show by cross multiplication of the numerators and the denominators that  $s^{ol} < s^{mp}$ , because  $-2n(n - 1)\gamma\hat{a} > 0$ . It means that the stock of pollution in steady state is higher in the linear Markov-perfect equilibrium than in the open-loop equilibrium. The intuition is that in the linear Markov-perfect equilibrium, each country emits more arguing that an extra unit of emissions will be partly offset by the other

countries, because the other countries observe a higher stock of pollution and react with lower emissions.

For some time, it was the common understanding that the Markov-perfect equilibrium is worse than the open-loop equilibrium. The idea was that a commitment to open-loop strategies is beneficial but in the absence of a commitment device, the Markov-perfect equilibrium is more realistic. However, Tsutsui and Mino (1990) showed, for another linear-quadratic differential game, that non-linear Markov-perfect equilibria exist, with the steady state close to the cooperative steady state. The other differential game was a dynamic duopoly with sticky prices (Fershtman and Kamien 1987). Dockner and Long (1993) showed the result for the game of international pollution control. The procedure is to manipulate the dynamic programming Eq. (20), where the optimal controls  $y_i(s)$  and the values functions  $V_i(s)$  are the same for each country  $i$ . When dropping the subscript  $i$ , Eq. (20) becomes

$$rV = \beta y - 0.5y^2 - 0.5\gamma s^2 + V'(ny - \delta s), y = \beta + V'. \quad (25)$$

The first thing to note is that (25) is a differential equation in  $V(s)$  that does not have a boundary condition. This explains the multiplicity of solutions  $V(s)$ , and thus  $y(s)$ . Dockner and Long (1993) proceed as follows. In the first step,  $r = 0$  and  $V'$  is replaced by  $y - \beta$  so that  $y$  can be solved. In the second step, this solution for  $y$  (for  $r = 0$ ) is extended with a term that is found by differentiating (25), replacing  $V'$  by  $y - \beta$  in the left-hand side, and solving the resulting differential equation in the unknown term. This leads to a set of non-linear feedback controls  $y$ , and finally to a set of corresponding steady states with the condition  $ny = \delta s$ . The expressions are complicated but can be found in Dockner and Long (1993) for  $n = 2$ . They arrive at the important conclusion that by using non-linear strategies, the Markov-perfect steady state can be decreased to a level that is arbitrarily close to the cooperative steady state, if the discount rate  $r$  is sufficiently low. Note that this resembles the folk theorem in repeated games where a subgame perfect equilibrium can support the cooperative outcome if the discount rate  $r$  is sufficiently low. Section 2.6 returns to this issue.

The procedure described above becomes easier by replacing  $V'$  by  $y - \beta$  in Eq. (25), differentiating (25), and replacing  $V'$ , which yields a differential equation in  $y$ :

$$[(2n - 1)y - \delta s - (n - 1)\beta]y' = (r + \delta)y + \gamma s - (r + \delta)\beta. \quad (26)$$

The solutions  $y(s)$  of (26) are of course the same as in Dockner and Long (1993). The linear  $y(s) = 2\hat{a}s + \hat{b} + \beta$ , where  $\hat{a}$  and  $\hat{b}$  are given by (21), is a solution of (26). Note that since  $y - \beta = V'$ ,  $y' = V''$ , Eq. (26) reduces to  $(y - \delta s)V'' = (r + \delta)V' + \gamma s$  for  $n = 1$ , which is the familiar differential equation  $\dot{\lambda}(t) = (r + \delta)\lambda(t) + \gamma s(t)$  for the co-state  $\lambda = V'$  in the time domain. For  $n > 1$  it is also possible to rewrite Eq. (26) as a system of two differential equations but of course, the underlying variable cannot be time because the term between brackets in front of  $y'$  is not the state transition. Section 2.6 applies this method, which was introduced by Dockner and Wagener (2014) for this type of problems.

### 2.6 Comparing the Equilibria

Because of symmetry, the subscript  $i$  is dropped in the equilibrium controls  $y_i$ , as in the last part of the previous section. For  $n = 2$ , the necessary conditions for the open-loop equilibrium can be rewritten from Eq. (16) as the dynamical system

$$\begin{aligned} \dot{s}(t) &= 2y(t) - \delta s(t), s(0) = s_0, \\ \dot{y}(t) &= \gamma s(t) + (r + \delta)(y(t) - \beta), \end{aligned} \tag{27}$$

and the necessary conditions for the cooperative solution can be rewritten from Eq. (18) as the dynamical system

$$\begin{aligned} \dot{s}(t) &= 2y(t) - \delta s(t), s(0) = s_0, \\ \dot{y}(t) &= n\gamma s(t) + (r + \delta)(y(t) - \beta). \end{aligned} \tag{28}$$

In both cases, the solution simply follows the stable line  $y(s)$  towards the steady states, which are given by Eqs. (17) and (19), respectively.

As mentioned at the end of the previous section, condition (26) for the Markov-perfect equilibrium can be rewritten as the dynamical system

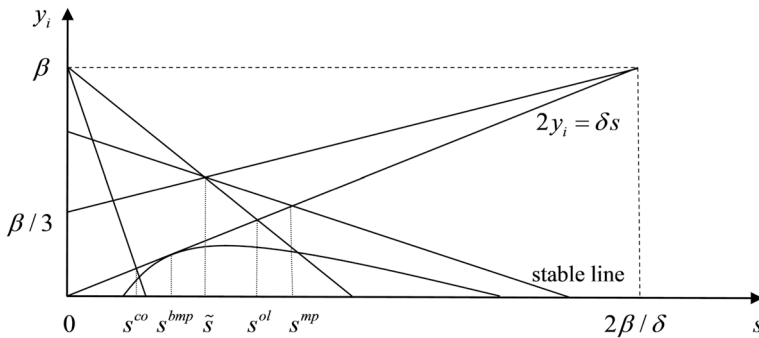
$$\begin{aligned} \dot{s}(u) &= 3y(u) - \beta - \delta s(u), s(0) = s_0, \\ \dot{y}(u) &= \gamma s(u) + (r + \delta)(y(u) - \beta), \end{aligned} \tag{29}$$

but note that the variable  $u$  cannot be time because the state transition would not be correct. However, the system (29) provides the candidates  $y(s)$  for the Markov-perfect equilibrium. The intersections of these candidates with the steady-state line  $2y(s) = \delta s$  provide the steady states, which can be stable or not stable. One candidate is the linear Markov-perfect equilibrium that coincides with the stable line of the dynamical system (29). The slope of this stable line is the direction of the stable eigenvector of the matrix with the rows  $[-\delta \ 3]$  and  $[\gamma \ r + \delta]$ . The characteristic polynomial of this matrix is  $\mu^2 - r\mu - \delta(r + \delta) - 3\gamma$ , so that the negative root or the negative eigenvalue is equal to  $\mu = 0.5\left(r - \sqrt{(r + 2\delta)^2 + 12\gamma}\right)$ . It follows that a stable eigenvector is  $[1 \ 2\hat{a}]'$ , where  $\hat{a}$  is given by Eq. (21) for  $n = 2$ . The stable line becomes  $y(s) = 2\hat{a}s + k$  where  $k$  is such that this line passes through the steady state of the system (29) given by

$$\tilde{s} = \frac{2\beta(r + \delta)}{3\gamma + \delta(r + \delta)}, \tilde{y} = \frac{\beta(\gamma + \delta(r + \delta))}{3\gamma + \delta(r + \delta)}. \tag{30}$$

It follows that  $k = \tilde{y} - 2\hat{a}\tilde{s}$ . Section 2.5 derives the linear Markov-perfect equilibrium  $y(s) = 2\hat{a}s + \beta + \hat{b}$  with  $\beta + \hat{b} = \beta(r + \delta - 2\hat{a})/(r + \delta - 6\hat{a})$  according to Eq. (21) for  $n = 2$ . Using that  $\hat{a}$  is the root of  $6\hat{a}^2 - (r + 2\delta)\hat{a} - 0.5\gamma = 0$ , it is straightforward to show that  $\tilde{y} - 2\hat{a}\tilde{s} = \beta + \hat{b}$ , so that the linear Markov-perfect equilibrium coincides with the stable line of the dynamical system (29), which is depicted in Fig. 1. This line intersects the steady-state line  $2y(s) = \delta s$  in  $s^{mp}$  (Eq. (23) for  $n = 2$ ).

However, the system (29) provides many more candidates  $y(s)$  for the Markov-perfect equilibrium. Figure 1 shows the solution curve  $y(s)$  that is tangent to the steady-state line  $2y(s) = \delta s$ . The solution curves in between this curve and the linear one intersect the steady-state line from above to below, so that these steady states are stable. Figure 1 also depicts the phase diagrams and the steady states  $s^{ol}$  and  $s^{co}$  for the open-loop equilibrium and cooperative solution, respectively. The solution curve that is tangent to the steady-state



**Fig. 1** Comparing equilibria

line represents the best Markov-perfect equilibrium in the sense that the steady state  $s^{bmp}$  is the closest to the cooperative steady state  $s^{co}$ . It can be shown that the steady state  $s^{bmp}$  converges to the cooperative steady state  $s^{co}$  if the discount rate  $r$  goes to 0. This is the main result of Dockner and Long (1993). Rubio and Casino (2002) put forward that this only holds in a certain range for the initial stock of pollution  $s_0$ . Indeed, steady state  $s^{bmp}$  is only stable from the right and not from the left, so that it cannot be reached from  $s_0 < s^{bmp}$ . Furthermore,  $s^{bmp}$  can also not be reached from  $s_0$  that are larger than the intersection of the solution curve with the  $s$ -axis. Dockner and Wagener (2014) suggest a way out by allowing a discontinuity in the solution curves. Section 3.4 returns to this issue.

### 3 Differential Games in Non-linear Systems with Tipping Points

A new challenge arises when the state transition is more complicated. Some curvature in the function for state transition does not essentially change the analysis, but the new challenge arises when the state transition is convex-concave so that the concavity of the decision problem is lost. This became important with the rise of ecological economics, since ecological systems with tipping points have this type of dynamics (Scheffer et al. 2001). Much earlier, the first optimal control analysis for this type of problem already appeared in the economics literature (Skiba 1978). At the time, this paper on optimal growth drew little attention, because the assumption of a convex-concave production function was not considered to be realistic. However, this type of state transition is very relevant in ecological systems. A typical result in the optimal control of these systems is the occurrence of a sequence of alternating stable and unstable steady states (Brock and Starrett 2003). In case of the prototype model of a lake, for example, a saddle point stable steady state appears in both the clean and the polluted region of the lake, with an unstable steady state in between (Mäler et al. 2003). The maximum principle provides necessary conditions in the form of candidate solution trajectories towards the stable steady states. For a large range of initial conditions of the lake, several trajectories exist leading to one or the other stable steady state. The optimal trajectory is the one with the highest value for the objective. For one initial condition, two optimal trajectories occur, with the same value, each moving to another stable steady state. This indifference point is called a *Skiba point*: it divides the set of possible initial conditions of the lake in two areas, each having one of the stable steady states as the attractor. If the lake is initially relatively clean, it is optimal to stay in the clean area

and move to the clean steady state. However, if the lake is initially already polluted, it is optimal stay in the polluted area and move to the polluted steady state.

The lake problem becomes a differential game in the case of more users of this public resource. The open-loop analysis of the symmetric differential game with more users produces similar dynamics as the optimal control analysis with one user. A combination of more users and a higher weight for the damage to the lake in the objective function yields the same prototype dynamics, with a saddle-point stable steady state in the clean and the polluted area of the lake and an unstable steady state in between. However, this problem does not have a Skiba point, because the differential game does not have one objective that discriminates between the different possible trajectories. In the game, the set of initial conditions splits into three areas. Two areas have one equilibrium with one of the stable steady states as attractor, and a third area in the middle has two equilibria, moving to one or the other stable steady state. Because the game is symmetric, the users can coordinate in this third area on the equilibrium with the highest value for all users, which is like a Skiba point but it is not the same.

The users of the lake can also cooperate and solve the optimal control problem with the higher weight for the damage to the lake in the objective function. If this weight is high enough, the lake model loses the unstable steady state and the stable steady state in the polluted area of the lake. It follows that for each initial condition, the optimal trajectory moves to the stable steady state in the clean area of the lake. This implies that if the lake is heavily polluted, the benefits of cooperation are large, because the benefits of being able to move to the clean area of the lake are added to the usual benefits of cooperation. A high level of pollution gives a high incentive to cooperate. After reaching the clean area of the lake, the benefits of cooperation become smaller and cooperation may break down, but the resulting equilibrium remains in the clean area of the lake.

The Markov-perfect or feedback equilibria of this differential game are, of course, non-linear, and the same type of multiplicity arises as for non-linear equilibria in the linear-quadratic differential game (Kossioris et al. 2008). The steady state of the best Markov-perfect equilibrium lies close to the steady state of the cooperative outcome, regardless of whether the initial condition is in the clean or in the polluted area. However, if the initial condition is such that the open-loop equilibrium can reach its clean steady state, the best Markov-perfect equilibrium does not necessarily have a higher value than the open-loop equilibrium, because the equilibrium trajectory proves to be less favourable. However, the Markov-perfect equilibrium has two clear advantages: the users cannot be trapped in the polluted area of the lake, as may happen in the open-loop equilibrium, and the steady state lies close to the cooperative steady state.

### 3.1 The Lake Model

Limnologists study lakes and observed that the state of a lake can suddenly change from a healthy state to an unhealthy state, with a big loss of ecosystem services such as fish, clean water, and amenities. The reason is the accumulation of phosphorus in the lake, which is released on the lake from neighbouring agriculture. This accumulation must be non-linear to explain the occurrence of these tipping points. A model for a lake that explains these observations uses what is called a Holling type III functional response (Carpenter and Cottingham 1997; Scheffer 1997). In a simple reduced form, the model for a lake becomes the differential equation

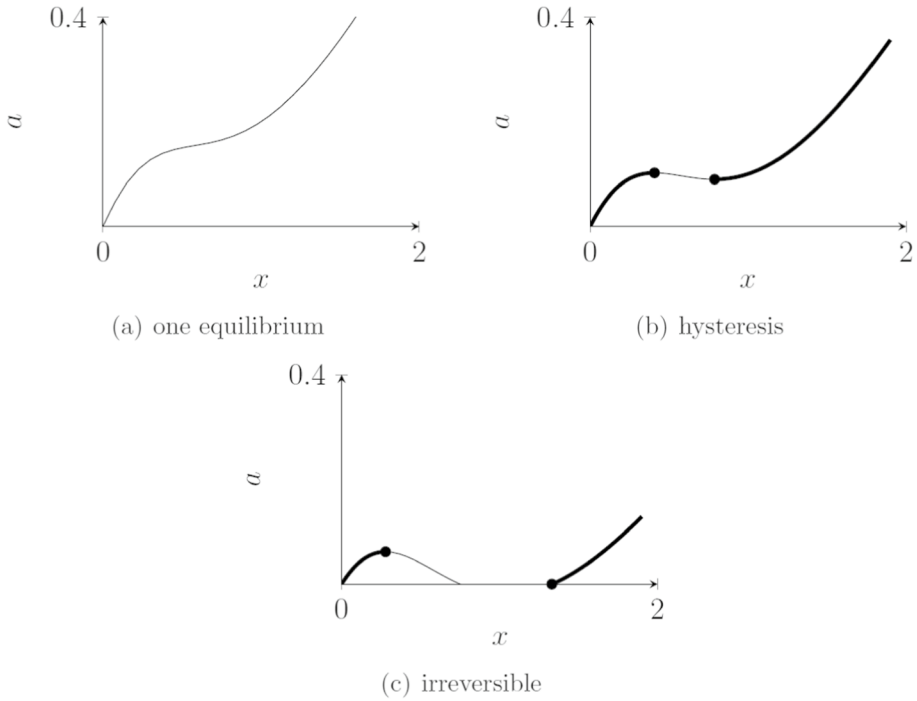


Fig. 2 Lake steady states

$$\dot{x}(t) = a(t) - f(x(t)), f(x) = bx - \frac{x^2}{1 + x^2}, x(0) = x_0, \tag{31}$$

where  $x$  represents the stock of phosphorus in the lake,  $a$  the loading of phosphorus on the lake, and  $b$  the natural outflow of phosphorus from the lake. The non-linear term results from interaction with sediments on the bottom of the lake. The full lake model contains a slowly changing variable to capture this interaction. The end of this section returns to this extension. The full lake model also contains an equation for the loading of phosphorus  $a$ , based on estimations with past data. In the sequel, however,  $a$  is a control variable, with the aim to maximize the benefits of agriculture, represented by the level of loading  $a$ , corrected for the damage of the stock of pollution  $x$ .

Figure 2 presents three cases for the steady states  $a = f(x)$  of the lake, for a small, an intermediate ( $0.5 < b < 0.65$ ), and a larger value of the parameter  $b$ .

The curve in Fig. 2a is non-linear, but the curvature is not strong, so that each  $a$  has one equilibrium  $x$ , and the analysis essentially does not change. Figures 2b and 2c have tipping points. Starting in 0 and increasing the loading  $a$  yields low equilibrium values of  $x$  first but at the left black dot, the equilibrium value of  $x$  jumps up. It is not possible to jump back immediately, but the value of  $a$  must be decreased to the right black dot to lower the equilibrium value of  $x$  again (a hysteresis effect). This process is clearer in Fig. 2c but in that figure, it is not possible to return to low equilibrium values of  $x$  (irreversibility), because  $a \geq 0$ . Figure 2b is the prototype lake model that will be used in the sequel.



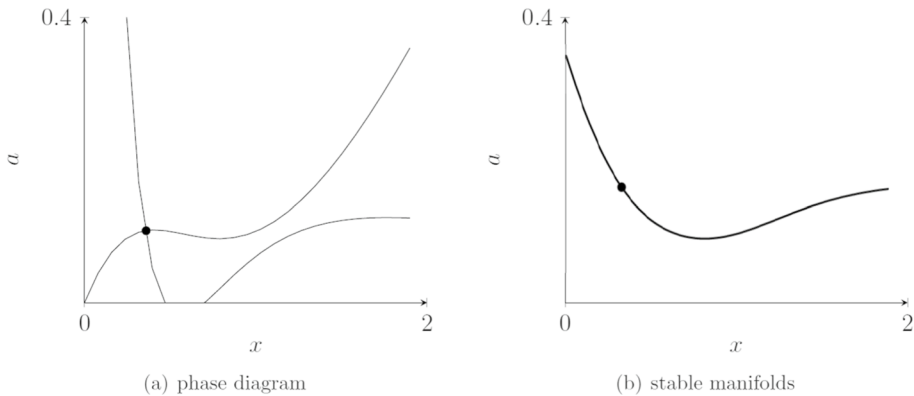


Fig. 3 Optimal control solution,  $c = 1$

### 3.2 Optimal Control Solution

The optimal control problem becomes (Mäler et al. 2003)

$$\max_{a(\cdot)} \int_0^{\infty} e^{-rt} [\ln a(t) - cx^2(t)] dt, \tag{32}$$

subject to (31). The logarithmic benefit function is convenient, because the cooperative solution with  $n$  agents is the same as the optimal control solution (see below).

The Hamiltonian  $H(x, a, \lambda) = \ln a - cx^2 + \lambda(a - f(x))$  yields the optimality condition  $1/a = -\lambda$  in (3), and then the dynamical system consisting of (31) and

$$\dot{\lambda}(t) - r\lambda(t) = f'(x(t))\lambda(t) + 2cx(t). \tag{33}$$

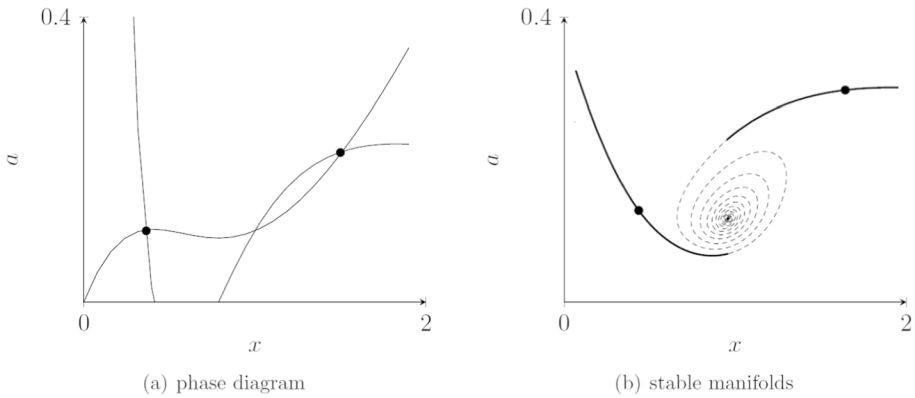
Using  $\dot{\lambda} = \dot{a}/a^2$ , Eq. (33) becomes a differential equation in the control  $a$ , i.e.

$$\dot{a}(t) = -(r + f'(x(t))a(t) + 2cx(t)a^2(t), f'(x) = b - \frac{2x}{(1 + x^2)^2}. \tag{34}$$

Figure 3a presents the steady states  $a = f(x)$  and  $a = (r + f'(x))/2cx$  for Eqs. (31) and (34), respectively, for  $b = 0.6$ ,  $c = 1$ , and  $r = 0.03$ . For these values of  $c$  and  $r$ , the two steady-state curves intersect only once, close to the tipping point as indicated in Fig. 2b. These curves divide the  $(x, a)$ -plane into areas where the variables  $x$  and  $a$  either increase or decrease. This is the phase diagram.

In this case, the analysis is not essentially different from the analysis in the linear case in Sect. 2. The dynamical system consisting of Eqs. (31) and (34) has only one initial condition, and therefore many possible solutions, but the optimal solution must be the stable manifold that converges towards the steady state of the dynamical system, indicated by the black dot. Figure 3b presents the trajectory of the optimal solution. The starting point on the stable manifold is given by the initial condition  $x(0) = x_0$ .

Figure 4 presents the analysis if the value of  $c$  is decreased to  $c = 0.5$ . This means that the relative weight of the damage to the lake  $cx^2$  with respect to the benefits  $\ln a$  is lower. A similar effect occurs if the discount rate  $r$  is higher.



**Fig. 4** Optimal control solution,  $c = 0.5$ , Skiba point

Figure 4a shows that the dynamical system has three steady states now. The two steady states, indicated by black dots, lie on a stable manifold, and thus are saddle-point stable. In between these steady states lies an unstable steady state. This alternating pattern of stable and unstable steady states is typical in these type of problems (Brock and Starrett 2003). Figure 4b presents the stable manifolds that curl out of the unstable steady state. There is a range of initial conditions  $x_0$  with an infinite choice of stable manifolds that either move to the left or to the right stable steady state. It is obviously better to start on the upper or the lower curve than to start somewhere inside. Furthermore, the optimal value of the objective determines whether it is better to start on the upper or the lower curve. For one initial condition  $x_0^s$ , the values of moving to the left or to the right stable steady state are the same. This is the Skiba point. It has the property that if  $x_0 \leq x_0^s$ , the optimal trajectory ends up in a clean steady state and if  $x_0 \geq x_0^s$ , the optimal trajectory ends up in a polluted steady state. This simply means that if the lake is already polluted and if the relative weight of the damage to the lake is small, it is optimal to accept the damage to the lake and to enjoy the higher benefits. However, if the lake is not so very polluted yet, it is optimal to clean it up.

### 3.3 Open-Loop Equilibrium

In the case of  $n$  users of the ecosystem services of the lake, who all release phosphorus on the lake, the differential game becomes (Mäler et al. 2003)

$$\max_{a_i(\cdot)} \int_0^\infty e^{-rt} [\ln a_i(t) - cx^2(t)] dt, i = 1, 2, \dots, n, \quad \text{subject to} \quad (35)$$

$$\dot{x}(t) = \sum_{i=1}^n a_i(t) - f(x(t)), f(x) = bx - \frac{x^2}{1+x^2}, x(0) = x_0. \quad (36)$$

The Hamiltonians  $H_i(x, a_i, \lambda_i) = \ln a_i - cx^2 + \lambda_i(a_i + \sum_{j \neq i}^n a_j - f(x))$  yield the optimality conditions  $1/a_i = -\lambda_i$ , and then the dynamical system consisting of (36) and

$$\dot{\lambda}_i(t) - r\lambda_i(t) = f'(x(t))\lambda_i(t) + 2cx(t), i = 1, 2, \dots, n. \tag{37}$$

The co-states  $\lambda_i$  and the loadings  $a_i$  are the same for  $i = 1, 2, \dots, n$ . Using  $\dot{\lambda}_i = \dot{a}_i/a_i^2$ , Eq. (37) becomes a differential equation in the total loading  $a$ , i.e.

$$\dot{a}(t) = -(r + f'(x(t))a(t) + 2\frac{c}{n}x(t)a^2(t)), a = \sum_{i=1}^n a_i, f'(x) = b - \frac{2x}{(1+x^2)^2}. \tag{38}$$

The dynamical system consisting of (36) and (38) is the same as the dynamical system consisting of (31) and (34), except for the additional parameter  $n$ , which indicates the number of users of the lake. This implies that in case  $c = 1$  and  $n = 2$ , Fig. 4 provides the stable manifolds for the trajectories of the open-loop equilibrium. However, in this case there is no Skiba point because there is no objective that differentiates between the possible trajectories. For each initial condition  $x_0$  in the range with the choice of stable manifold, the open-loop equilibrium is not unique. Both trajectories (along the upper curve towards the right stable steady state and along the lower curve towards the left stable steady state) are trajectories of an open-loop equilibrium (Grass et al. 2017). In this symmetric game, the equilibrium values of the objectives are the same for the two users. Therefore, in this case, the Skiba point still has a meaning in the sense that the users can coordinate on the best open-loop equilibrium, and move to the clean steady state, if  $x_0 \leq x_0^s$ , and to the polluted steady state, if  $x_0 \geq x_0^s$ .

The choice for the logarithmic benefit function means that if the  $n$  users cooperate, the dynamical system (31) and (34) for the optimal control solution results. Indeed, with the Hamiltonian  $H(x, a_1, \dots, a_n, \lambda) = \sum_{i=1}^n \ln a_i - ncx^2 + \lambda(\sum_{i=1}^n a_i - f(x))$ , the optimality conditions are  $1/a_i = -\lambda$ , leading to the dynamical system (31) and (34). Since  $c = 1$ , Fig. 3 provides the stable manifold for the cooperative solution.

If the initial condition in Fig. 4 lies to the right of the curls, the open-loop equilibrium is trapped in the polluted area of the lake. In that case cooperation gets the users out of the polluted area, because then they will choose the stable manifold in Fig. 3. The benefits of cooperation are high now. However, if the initial condition in Fig. 4 lies to the left of these curls, the steady state of the open-loop equilibrium lies close to the cooperative steady state. The numbers for the stock of phosphorus are 0.393 and 0.353, respectively. These numbers are just to the left of the tipping point 0.408.

### 3.4 Markov-Perfect Equilibria

Using Eq. (7), the dynamic programming equations for the lake differential game in the previous section are (Kossioris et al. 2008)

$$rV_i(x) = \max_{a_i} [\ln a_i - cx^2 + V'_i(x)(a_i + \sum_{j \neq i}^n a_j - f(x))], i = 1, 2, \dots, n. \tag{39}$$

The value functions  $V_i$  and the loadings  $a_i$  are the same for  $i = 1, 2, \dots, n$ . Replacing  $V'_i$  by  $-1/a_i$  (using the optimality condition), differentiating Eq. (39), replacing  $V'_i$  by  $-1/a_i$  again, and rearranging yields the differential equation in  $a_i$

$$[(a_i - f(x))a'_i = -(r + f'(x))a_i + 2cxa_i^2. \tag{40}$$

Note that if there is only one user, so that  $a_i$  denotes total loading  $a$ , Eq. (40) can be replaced by Eqs. (31) and (34) in the time domain. Equation (40) does not have a boundary condition but for the equilibrium trajectories that converge to a steady state  $x_{ss}$ , the boundary condition  $a_i(x_{ss}) = f(x_{ss})/n$  can be used.

Equation (40) is an Abel differential equation of the second kind (Murphy 1960). The Markov-perfect equilibrium curves  $a_i(x)$  were found by using the ode solver *ode15s* of Matlab. Kossioris et al. (2008) present these curves for the parameter values  $n = 2$ ,  $b = 0.6$ ,  $c = 1$ , and  $r = 0.03$ , and the result is very similar to what was found in Sect. 2.5. Starting at a high initial condition  $x_0$ , many curves can be chosen that intersect the individual steady-state curve  $f(x)/2$ , which is half the curve in Fig. 2b. When the equilibrium curves intersect the individual steady-state curve from above to below, the resulting steady states are stable. The Markov-perfect steady state that is closest to the cooperative steady state arises when the curve  $a_i(x)$  is below and tangent to the steady-state curve  $f(x)/2$ . The number for the stock of phosphorus in this case is 0.38, and this number approaches the number for the cooperative steady state (0.353) in case the discount rate  $r$  goes to 0. Unfortunately, however, the value of the objective in this best Markov-perfect equilibrium is only close to the cooperative optimal value, if the initial condition  $x_0$  lies close to the steady state. Otherwise, the value is much worse, so that then the Markov-perfect equilibrium is not a good alternative for cooperation. The main advantage of the Markov-perfect equilibrium is that it gets the users out of the polluted area of the lake without the need for cooperation.

Starting at an initial condition  $x_0 < 0.38$ , the same issues arise as in Sect. 2.5. If the initial condition  $x_0 < 0.17$ , it is not possible to reach the steady state 0.38. However, if  $0.17 < x_0 < 0.38$ , it is possible to reach 0.38 in two steps by first moving beyond 0.38, following a trajectory above the steady state curve, and then jumping down to the best Markov-perfect equilibrium. Dockner and Wagener (2014) provide another solution to this problem. They present a Markov-perfect equilibrium that is discontinuous but that is well-defined for lower stocks of pollution. They use a different methodology. As was noted in Sect. 2.5 for Eq. (26), Eq. (40) can be rewritten as a system of two differential equations,  $\dot{x} = a_i - f(x)$  and  $\dot{a}_i = -(r + f'(x))a_i + 2cxa_i^2$ , but the underlying variable cannot be time, of course. In this way, they manage to derive the discontinuous Markov-perfect equilibrium that is defined for all stocks of pollution.

### 3.5 The Full Lake Model

The simple reduced form (31) for the lake model ignores the slowly changing stock of phosphorus in the sediment or the “mud” of the lake, and the interaction between the stocks of phosphorus in the water and in the mud of the lake (Carpenter 2005). Grass et al. (2017) present the optimal control solution and the open-loop equilibrium for the extended lake model. On the dimension of the fast-changing stock of phosphorus in the water, the dynamics are very similar to what is found above. The slowly changing stock of phosphorus in the mud mainly adjusts the dynamical system to the final steady state. It is beyond the scope of this paper to present these results here. Grass et al. (2017) use the toolbox OCMat,<sup>1</sup> which is especially designed for this type of problems and uses the Matlab solvers. This is already a serious computational effort. The Markov-perfect equilibrium would be even more complicated and is still open for research. One may wonder if the additional

<sup>1</sup> [http://orcos.tuwien.ac.at/research/ocmat\\_software](http://orcos.tuwien.ac.at/research/ocmat_software).

effort will lead to more insights. On the other hand, software for this type of problems will become available soon, and then it will be relatively easy to perform this type of analyses and to see what it provides.

#### 4 Stackelberg Equilibria and Time-Inconsistency in Macroeconomics

Besides investigating Nash equilibria, the theory of differential games also investigated Stackelberg equilibria (see Simaan and Cruz 1973a; b). In a Nash game, the players decide simultaneously and in a Stackelberg game, the players decide sequentially, in a leader–follower structure. In a Nash game, each player formulates a best response to the strategy of the other players, and a Nash equilibrium requires the best responses to hold simultaneously. In a Stackelberg game with two players, the follower formulates a best response to the strategy of the leader, and the leader optimizes while taking account of the best response of the follower. If the information structure is open loop, the players formulate a time path of actions. This implies that the best response of the follower in early stages is also a reaction to the strategy of the leader in later stages. However, after some time has passed, this effect is gone, and the leader and the follower will want to change their strategies. This property of the open-loop Stackelberg equilibrium is called *time-inconsistency*: reconsidering the game after some time has passed will change the outcome. By construction, the feedback or Markov-perfect Stackelberg equilibrium is time-consistent. This property of this Stackelberg equilibrium is even stronger because the outcome of the game does not change at any time, but also does not change for any value of the state. Therefore, feedback or Markov-perfect is also called strongly time-consistent. Weakly time-consistent then refers to time-consistency *on* but not *off* the equilibrium trajectory. It is interesting to note here that the open-loop Nash equilibrium proves to be weakly time-consistent but not strongly time-consistent. Requiring strong time-consistency for the open-loop Nash equilibrium means switching to the feedback or Markov-perfect Nash equilibrium.

Time-inconsistency of the open-loop Stackelberg equilibrium has become an important concept in macroeconomics. The government has the incentive to change an optimal policy after some time has passed, and rational agents will expect this to happen, so that the optimal policy will fall apart (Kydland and Prescott 1977). The conclusion was that optimal control theory is not applicable to economic planning when expectations are rational. However, it is more precise to say that this is not an optimal control problem but a differential game. The open-loop Stackelberg equilibrium with the government as leader and the economic agents as followers is time-inconsistent. The issue is solved by imposing a time-consistency constraint. The strong form, with time-consistency on and off the equilibrium trajectory, requires basing policy on the feedback or Markov-perfect Stackelberg equilibrium. However, the government can do better if it can commit to its open-loop Stackelberg strategy. This has set the stage for discussions on rules versus discretion in macroeconomics, but the same discussion applies to government policy in environmental economics of course. The introduction of the issue of time-inconsistency in macroeconomics was considered so important that the Prize in Economic Sciences in memory of Alfred Nobel was awarded to Kydland and Prescott in 2004, for the paper in 1977. The next section provides more details of this analysis.

#### 4.1 The Kydland-Prescott Model

Kydland and Prescott (1977) argue that an optimal plan is time inconsistent. Suppose there are two periods. The social objective function is  $S(\pi_1, \pi_2, x_1, x_2)$ , where  $(\pi_1, \pi_2)$  are the policies in periods 1 and 2, and  $(x_1, x_2)$  are the decisions of the economic agents. These decisions are functions of the whole policy  $(\pi_1, \pi_2)$  and the past decisions, i.e.  $x_1 = X_1(\pi_1, \pi_2)$  and  $x_2 = X_2(x_1, \pi_1, \pi_2)$ . The optimal plan  $\pi_2$  must satisfy

$$\frac{\partial S}{\partial \pi_2} + \frac{\partial S}{\partial x_1} \frac{\partial X_1}{\partial \pi_2} + \frac{\partial S}{\partial x_2} \left( \frac{\partial X_2}{\partial x_1} \frac{\partial X_1}{\partial \pi_2} + \frac{\partial X_2}{\partial \pi_2} \right) = 0. \quad (41)$$

When period 1 has passed, the optimal policy  $\pi_2$  in period 2 must satisfy

$$\frac{\partial S}{\partial \pi_2} + \frac{\partial S}{\partial x_2} \frac{\partial X_2}{\partial \pi_2} = 0. \quad (42)$$

Only if the effect of  $\pi_2$  on  $x_1$  is zero or if the direct and indirect effect of  $x_1$  on  $S$  are zero, the conditions (41) and (42) are the same. Otherwise, the optimal plan  $\pi_2$  is time-inconsistent in the sense that the optimal policy  $\pi_2$  in period 2 will change when period 1 has passed and the policy  $\pi_2$  is reconsidered.

This is precisely the time-inconsistency of the open-loop Stackelberg equilibrium. The policy maker is the leader, and the economic agents are the followers. The optimal plan results from maximizing the social objective function while taking the best response of the followers into account. This yields the condition (41). However, when period 1 has passed, the optimal policy changes and satisfies condition (42).

Kydland and Prescott (1977) argue that if the economic agents have some knowledge of how the future policies will change, the optimal plan will not work. They conclude that optimal control theory is not the appropriate tool for economic planning. However, it is better to say that the economic system is not an optimal control problem, but it is a Stackelberg differential game. Then the conclusion is that the policy of the leader must be time-consistent, in case of some form of rational expectations. This can be achieved by determining the policy by backward induction. This means that Eq. (42) must hold for  $\pi_2$ , and then  $\pi_1$  can be determined given the outcome for  $\pi_2$  as a function of the past policies and decisions of the economic agents. It is interesting to note that time-inconsistency of Stackelberg equilibria was already an issue in differential game theory in the early seventies, but it took some time before this became an important issue in macroeconomics. The example shows that differential game theory is crucial in the case of intertemporal strategic interaction, which is core to economics.

## 5 Conclusion

The purpose of this paper is to carefully explain the methodology of differential games, and to present some characteristic applications. The material is mostly available in the literature but may not always be easily accessible. Differential games are important for economics, because it is the natural framework for a situation with strategic interaction between economic agents and with stock-flow dynamics of important variables such as the stock of resources or pollution.

A simple example from over 50 years ago shows the several natural ways of solving a dynamical game, and the generalizations to the solution techniques for optimal control problems. The maximum principle and dynamic programming yield the same result for optimal control problems but yield different equilibria when used to solve differential games. The maximum principle, with strategies that are only a function of time, yields the open-loop Nash equilibrium, whereas dynamic programming yields the feedback or Markov-perfect Nash equilibrium, where the strategies are a function of time and state (the stock) and where subgame perfectness holds.

The game of international pollution control is a linear-quadratic differential game, with a quadratic objective and a linear state transition. The benchmark is full cooperation, which is an optimal control problem. The steady state of the open-loop equilibrium is closer to the cooperative steady state than the steady state of the linear Markov-perfect equilibrium, but non-linear Markov-perfect equilibria exist and the best Markov-perfect equilibrium has a steady state close to the cooperative one.

The lake game has a non-linear state transition with tipping points. In the case of two stable steady states, one in the clean and one in the polluted area of the lake, the optimal control problem has a Skiba point, which divides the initial conditions into an area with the clean steady state as attractor and an area with the polluted steady state as attractor. The many open-loop equilibria in such a case move either to the clean or the polluted steady state but for the initial conditions above a certain level, the game ends up in the polluted area. Cooperation of the users implies that the damage to the lake for the other users is taken into account, so that it can become optimal to move to a clean steady state from any initial condition. This means that cooperation gets the users out of the polluted area. The Markov-perfect equilibria get the users out of the polluted area as well and similar as in the other game above, the best Markov-perfect equilibrium moves to a steady state close to the cooperative one, although the welfare is much lower.

The last section of the paper turns to the concept of time-consistency in policy making. With the policy maker as the leader and economic agents as the followers, the setting is a Stackelberg differential game. The theory of differential games shows that the open-loop Stackelberg equilibrium is time-inconsistent, which means that the policy maker will change the policy after time has passed by. This result is relevant for all policy in economics, but it especially affected macroeconomics. It was argued that a policy plan will not work if the policy is time-inconsistent and if the economic agents have rational expectations. In 2004, this idea was awarded the Prize in Economic Sciences in memory of Alfred Nobel.

**Acknowledgement** I am grateful to Tamer Başar and Geert Jan Olsder for teaching me differential games, to Rick van der Ploeg, Anastasios Xepapadeas and Florian Wagener for their cooperation in applying these ideas, and to anonymous reviewers who helped to improve this paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Başar T, Olsder G-J (1982) *Dynamic noncooperative game theory*. Academic Press, New York
- Bellman R (1957) *Dynamic programming*. Princeton University Press, Princeton
- Brock WA, Starrett D (2003) Managing systems with non-convex positive feedback. *Environ Resour Econ* 26(4):575–602
- Carpenter SR (2005) Eutrophication of aquatic ecosystems: Bistability and soil phosphorus. *Proc Natl Acad Sci* 102(29):10002–10005
- Carpenter SR, Cottingham KL (1997) Resilience and restoration of lakes. *Conserv Ecol* 1:2
- Davis MHA (1977) *Linear estimation and stochastic control*. Chapman and Hall, London
- Dockner E, Van Long N (1993) International pollution control: cooperative versus noncooperative strategies. *J Environ Econ Manag* 25(1):13–29
- Dockner E, Wagener F (2014) Markov perfect Nash equilibria in models with a single capital stock. *Econ Theor* 56(3):585–625
- Dockner E, Jørgensen S, Van Long N, Sorger G (2000) *Differential games in economics and management science*. Cambridge University Press, Cambridge, UK
- Fershtman C, Kamien MI (1987) Dynamic duopolistic competition with sticky prices. *Econometrica* 55(5):1151–1164
- Fleming WH, Rishel RW (1975) *Deterministic and stochastic optimal control*. Springer-Verlag, Berlin
- Grass D, Xepapadeas A, de Zeeuw A (2017) Optimal management of ecosystem services with pollution traps: the lake model revisited. *J Assoc Environ Resour Econ* 4(4):1121–1154
- Isaacs, Rufus (1954–56). *Differential Games I, II, III, IV. Rand Cooperation Research Memorandum* RM-1391, 1399, 1411, 1468.
- Kossioris G, Plexousakis M, Xepapadeas A, de Zeeuw A, Mäler K-G (2008) Feedback Nash equilibria for non-linear differential games in pollution control. *J Econ Dyn Control* 32(4):1312–1331
- Kydland FE, Prescott EC (1977) Rules rather than discretion: the inconsistency of optimal plans. *J Polit Econ* 85(3):473–492
- Mäler K-G, Xepapadeas A, de Zeeuw A (2003) The economics of shallow lakes. *Environ Resour Econ* 26(4):603–624
- Murphy GM (1960) *Ordinary differential equations and their solutions*. Van Nostrand, Princeton
- Pontryagin LS, Boltyanskii VG, Gamkrelidze RV, Mishchenko EF (1962) *The mathematical theory of optimal processes*. Interscience, New York
- Rubio SJ, Casino B (2002) A note on cooperative versus non-cooperative strategies in international pollution control. *Resour Energy Econ* 24(3):251–261
- Scheffer M (1997) *Ecology of shallow lakes*. Chapman and Hall, London
- Scheffer M, Carpenter SR, Foley JA, Folke C, Walker B (2001) Catastrophic shifts in ecosystems. *Nature* 413:591–596
- Selten R (1975) Re-examination of the perfectness concept for equilibrium points in extensive games. *Int J Game Theory* 4:25–55
- Simaan M, Cruz JB Jr (1973a) On the Stackelberg strategy in nonzero-sum games. *J Optim Theory Appl* 11(5):533–555
- Simaan M, Cruz JB Jr (1973b) Additional aspects of the Stackelberg strategy in nonzero-sum games. *J Optim Theory Appl* 11(6):613–626
- Skiba A (1978) Optimal growth with a convex-concave production function. *Econometrica* 46(3):527–539
- Starr AW, Ho YC (1969a) Nonzero-sum differential games. *J Optim Theory Appl* 3(3):184–206
- Starr AW, Ho YC (1969b) Further properties of nonzero-sum differential games. *J Optim Theory Appl* 3(4):207–219
- Tsutsui S, Mino K (1990) Nonlinear strategies in dynamic duopolistic competition with sticky prices. *J Econ Theory* 52(1):131–161
- van der Ploeg F, de Zeeuw A (1992) International aspects of pollution control. *Environ Resour Econ* 2(2):117–139

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.