10. Games with Changing Payoffs

Reinoud Joosten, Hans Peters and Frank Thuijsman*

1. INTRODUCTION

A game with changing payoffs or actions is a dynamic game in which the payoffs or action sets may change from one decision moment to the next as a consequence of the actions played previously. Stochastic games, as well as differential games are, generally speaking, examples of such games. On the other hand, a repeated game is not a game with changing payoffs in this sense; although the payoffs may change over time, this is a consequence of, for instance, time discounting, and not of the actions played.

Our motivation for studying games with changing payoffs or action sets comes from the idea that by (not) performing certain actions the payoffs resulting from those actions may increase (decrease), or the set of available actions may change. Although this phenomenon may be called *learning* or *unlearning* (see Joosten *et al.* 1991), these expressions should be understood in a different way than is usual in the game-theoretic literature. By *(un)learning* we do not mean (un)learning how to play the game, nor gathering (or losing) information about the game. Rather, it should be interpreted as (un)learning how to perform a *physical* action – where *physical* can be taken in a broad sense. Let us clarify this by some examples.

In a dynamic duopoly situation a firm may choose to offer more than the Cournot–Nash equilibrium amount. The relative loss suffered may be compensated by enhancing its production technology – by the 'practical' production experience – or enlarging its market share. This is an example of a situation where (not) performing an action increases (decreases) the future payoffs resulting from that action.

An example from sports is the decathlon, where athletes may specialise in specific skills, not only depending on their own capabilities but in par-

* The authors thank Fernando Vega-Redondo and Thomas Ziesemer for comments and suggestions.

ticular also on the skills and specialisations of their adversaries. If a certain skill – say, high jumping – is stimulated or neglected, then future jumps will be higher or lower.

One might also think of countries competing in the world market, contemplating the adoption of new technologies. Learning-by-doing effects are to be anticipated on any adopted technology, whereas unlearning-by-not-doing effects must be anticipated on the 'traditional' activities. The experience of certain former colonies in sub-Saharan Africa (cf. Acharya 1981) may serve as an illustration in this context. Under colonialism intra-African economic ties were strongly discouraged in favour of economic ties with the coloniser. Economic structures and relationships within the colonies were transformed in the interests of the colonial power or in the interests of European settlers. African economic interests were generally disregarded, and African initiative was often heavily discouraged. For example in Kenya, Africans were prohibited from growing coffee until 1948–1949, and veterinary services for African-owned dairy cattle were withheld until 1955 (Heyer 1976). It is therefore not surprising that some of these former colonies found themselves at independence with little entrepreneurial and managerial know-how, an agricultural sector with little differentiation focused on production for the market of the coloniser, and an economy open to the coloniser, lacking important inter-industrial links and ties with neighbouring countries. The combined effects of not being able to 'learn' certain skills and processes fast enough in the post-colonial period to be competitive on the world market, and having 'unlearned' attractive alternatives which had been present in pre-colonial times, seem to have contributed to the problems which these former colonies face in industry and agriculture at present.

Such examples indicate that a variety of situations can be modelled as dynamic games with changing payoffs or changing actions. In particular, the choice a player may have between specialising on certain actions or trying to keep the spectrum of available and worthwhile actions as broad as possible, is an important feature of such games. Games like this have been analysed in the game-theoretic literature, mainly in the form of stochastic or differential games.

Before considering both types of games in somewhat more detail, a few words on the existing learning-by-doing models are in order. *Learning-by-doing* is the title of a pioneering paper by Arrow (1962). The existence of the possibility to learn by doing is not surprising. The novelty of learning-by-doing lies in its incorporation as a concept into economic theory. In a game-theoretic setting learning-by-doing is a different phenomenon, since learning-by-doing decisions also depend on what the other players do (see, in particular, the next section).

The purpose of this note is to present some examples of dynamic games with changing actions or payoffs. In section 2 we consider infinitely repeated matrix games where actions vanish if they have not been used for some time. Such games are a special type of stochastic game. In section 3 some differential games are analysed where each momentary action determines not only an immediate payoff but also influences a state variable which is part of the payoff function. Section 4 concludes the paper with a few remarks.

2. STOCHASTIC GAMES: VANISHING ACTIONS

A stochastic game (introduced by Shapley 1953) is characterised by a collection of states. In each state the players choose actions; these actions determine immediate payoffs as well as a probability distribution over the collection of states. The state at the next decision moment is determined on the basis of this probability distribution. The overall reward can be a discounted sum of immediate payoffs, or a limit of average payoffs; both criteria have been and are still being studied. A stochastic game clearly is an example of what we have called a dynamic game with changing actions.

Our first attempt to study (un)learning in the sense as described above, is Joosten *et al.* (1991). Two players repeatedly play a matrix game, where the entries of the matrix represent payoffs by the column player to the row player. Each player has a memory of a certain length, say r_1 and r_2 for players 1 and 2, respectively. If player 1, the row player, does not choose a certain row for r_1 consecutive times, then he loses the possibility to do so; that row is deleted from the matrix. Similarly for player 2, the column player, when he does not play a certain column for r_2 consecutive times. The payoff criterion is the limiting average payoff. Observe that this game is a stochastic game with a very special payoff/transition structure. The existence of limiting average ε-optimal strategies, for any $\varepsilon > 0$, follows from an established result in stochastic game theory (Mertens and Neyman 1981). The interesting aspect is that for some cases optimal strategies *can* be found and are relatively easy to describe. As an example, consider a 2 × 2 matrix game:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

If this game has a saddlepoint, as for instance in the specification then it is obvious that the one-shot optimal actions, namely the top row for

$$A = \begin{bmatrix} 1 & 2 \\ 0 & d \end{bmatrix},$$

player 1 and the left column for player 2, are, when repeatedly played, also optimal in the memory-restricted infinitely repeated game. In this case there is no proper 'learning-by-doing' or 'unlearning-by-not-doing'. The players concentrate on what already are their optimal strategies; by doing so, eventually they lose their suboptimal strategies, which only reinforces their incentives to play optimally – so to speak. This is similar to the one decisionmaker case, and a game-theoretic analysis sheds no further light on the situation.

The situation becomes different and more interesting if the original zero sum game does *not* have a saddlepoint, say $a \geq d > b \geq c$.

As an example, assume that both players have memory of length equal to 2. In this case, for player 1 it is optimal to start by playing each row with probability ½. If payoff a or d is realised, then he should play his second or first row, respectively, at the next stage, and keep switching rows as long as player 2 still has both columns available; as soon as player 2 loses a column, player 1 should play the payoff maximising row. If, at the first stage, payoff c or b is realised, then player 1 should play the first row forever. Player 2 has a similar optimal strategy. The expected payoff – the value of the game – is, thus, $\frac{1}{2}(b + d)$. In this game, both players at first keep both actions alive; actually, in optimal play their first moves are chance moves, and only from the second move on do the players play deterministically.

If both players have memories of length equal to three, then the optimal strategies are somewhat more complicated but can still be described. The value of the game is equal to $v := \frac{1}{4}(a + b + c + d)$ if this number is between b and d. It is equal to b if $v < b$, and it is equal to d if $v > d$. We refer to Joosten *et al.* (1991) for more details.

For an arbitrary but finite length of memory it is not easy to calculate or describe the optimal strategies in the above games. To some extent, this is due to the discrete nature of the game; the game is played in discrete time, and actions vanish suddenly. In the next section we consider a few examples of differential games with changing payoffs. In simple cases it is possible to calculate a certain type of Nash equilibrium by optimal control methods.

3. DIFFERENTIAL GAMES: CHANGING PAYOFFS

In a differential game, the players choose actions in continuous time, thereby receiving a flow of payoffs. Such actions are called *controls*, and they are chosen subject to certain constraints, in particular with respect to a *state variable*. Such a constraint is called the *state equation* or *transition equation*. The state plays a role similar to the state in a stochastic game; in the vanishing actions games of the preceding section, states are described by keeping track, for each possible action, of the number of times that action may not be played before it is lost. Differential games are often analysed by methods provided by optimal control theory, or by dynamic programming (see for instance Starr and Ho 1969a,b).

Differential games are used to model situations like common resource extraction. Suppose two countries use a common resource over a certain period of time. At each moment, their decisions to use an amount of the resource influence their profits (in a Cournot-like fashion), as well as the remaining stock of the resource (see, for instance, McMillan 1986). An additional assumption could be that prices might increase as the amount of the resource left for the future decreases. This would imply that the payoffs of the players change as a result of their previous actions.

3.1 An Investment Problem

In this subsection we analyse a differential game corresponding to a stylised economic problem of choosing between two ways to invest money. Specifically, we consider a two-player game in continuous time where at each point of time $t \in [0,\infty)$ each player has one (perfectly divisible) unit of money to invest. Each player can divide this one unit between on the one hand a project for which the payoff depends on the investments of both players, and on the other hand a project for which the payoff depends only on own investment. An investment in the first project will, moreover, result in an additional payoff stream, depending on both own investment and the investment of the opponent. This is meant to capture the idea of learning or unlearning as explained in the introduction. One may think of increasing or decreasing one's skill/technology[1] or market share. These additional payoffs constitute state variables.

Let $\alpha(t) \in [0,1]$ and $\beta(t) \in [0,1]$ denote the investment decisions at time t of players 1 and 2, respectively, in the first project. Let $g(\alpha(t), \beta(t))$ and $h(\alpha(t), \beta(t))$ denote the resulting immediate payoffs at time t for players 1 and 2, respectively. The function g can be assumed to have obvious properties, like being increasing in α and decreasing in β. Similarly for h. In this basic formulation, however, we do not need such assumptions. We just assume that both functions are continuously differentiable, but that assumption may also be relaxed. As functions of t, however, we require that α and β have only isolated points of discontinuity, in order to ensure the existence of the integrals below.

The immediate payoffs from investment in the second project at time t are equal to $1 - \alpha(t)$ and $1 - \beta(t)$, if $\alpha(t)$ and $\beta(t)$ are the investments in the first project, respectively. (Un)learning effects for player 1 are assumed to be captured by a state variable x depending on α as well as on β by the state equation $\dot{x}(t) = \alpha(t) - \beta(t)$ (where the dot denotes time derivative). The additional resulting payoff stream for player 1 is given by $x(t)e^{-rt}$, where r may be any real number. Here, $x(t)$ expresses the amount of 'learning' relative to the opponent, whereas e^{-rt} describes its long-run effect. Unlearning effects are stressed when r is positive; note that in that case in the long run the term $x(t)e^{-rt}$ practically vanishes, so that only short-term effects are interesting. A formulation of the problem where this is avoided, that is, where also long-term effects are interesting, is given in subsection 3.4.

Similarly, (un)learning effects for player 2 are given by a state variable y governed by the state equation $\dot{y}(t) = \beta(t) - \alpha(t)$. The corresponding additional payoff stream is given by $y(t)e^{-st}$, for some real number s. Note that $x(t) + y(t)$ is constant, so one can think of $x(t)$ as the market share of player 1 at time t. According to this interpretation, the constants x_0 and y_0 in the two maximisation problems to follow can be seen as the initial market shares, and it would be natural to choose $x_0 + y_0$ equal to 1. The case of actual learning would correspond to both initial values being set equal to 0.

We can now write down player 1's maximisation problem for any given investment plan $\beta(t)$ ($t \in [0,\infty)$) of player 2 and any discount factor ρ:

$$\begin{aligned} &\text{Maximise} \quad \int_0^\infty e^{-\rho t}\,[g(\alpha(t),\beta(t)) + (1 - \alpha(t)) + x(t)e^{-rt}]\,dt \\ &\text{subject to} \qquad \dot{x}(t) = \alpha(t) - \beta(t) \\ &\qquad\qquad\qquad x(0) = x_0, \alpha(t) \in [0,1]. \end{aligned} \tag{10.1}$$

Similarly, for player 2, given investments $\alpha(t)$ of player 1 at each moment $t \in [0,\infty)$:

$$\begin{aligned} &\text{Maximise} \quad \int_{\sim 0}^\infty e^{-\rho t}\,[h(\alpha(t), \beta(t)) + (1 - \beta(t)) + y(t)e^{-st}]\,dt \\ &\text{subject to} \qquad \dot{y}(t) = \beta(t) = \beta(t) - \alpha(t) \\ &\qquad\qquad\qquad y(0) = y_0, \beta(t) \in [0,1]. \end{aligned} \tag{10.2}$$

Thus, the players are assumed to maximise discounted streams of payoffs (with common discount factor ρ), given the investment plans of their opponents. The initial conditions for the state variables x and y are included to make the maximisation problems well-defined, but play no role in our analysis.

Observe that, in a seemingly more general but equivalent formulation, the terms $1 - \alpha$ and $1 - \beta$ in the objective functions could be taken into the functions g and h, respectively.

A simultaneous solution of problems (10.1) and (10.2) is a Nash equilibrium for this game. Depending on the nature of the strategies (investment plans) employed, we distinguish between open-loop strategies and closed-loop (feedback) strategies. In the latter case, strategies may depend on the state variables, and the players have the possibility to adapt their action choices while the game is being played. In the former case, a strategy depends only on time and not on the state variables. We will concentrate on open-loop strategies, which are much easier to calculate.

Solving problems (10.1) and (10.2) is a straightforward application of optimal control theory, specifically, of Pontryagin's maximum principle. The Hamiltonian corresponding to problem (10.1) is the function

$$H(\alpha,x,t,\lambda) = e^{-\rho t}\,[g(\alpha,\beta) + (1-\alpha) + xe^{-rt}] + \lambda[\alpha - \beta],$$

where the Lagrange multiplier (or costate variable) λ is also a function of t. Necessary conditions for a function α solving problem (10.1) are:

(a) At each t, α maximises[2] H. Thus, for an interior solution $0 < \alpha < 1$, we have $\partial H/\partial\alpha = 0$, hence

$$e^{-\rho t}[\partial g(\alpha,\beta)/\partial\alpha - 1] + \lambda = 0.$$

For a solution $\alpha = 0$ we have

$$e^{-\rho t}[\partial g(\alpha,\beta)/\partial\alpha - 1] + \lambda \leq 0,$$

and for a solution $\alpha = 1$ we have

$$e^{-\rho t}[\partial g(\alpha,\beta)/\partial\alpha - 1] + \lambda \geq 0.$$

(b) $\dot{x} = \partial H / \partial\lambda$, i.e. $\dot{x} = \alpha - \beta$.
(c) $\dot{\lambda} = -\partial H / \partial x$, i.e. $\dot{\lambda} = -\,e^{-(\rho+r)t}$.
(d) Transversality condition: $\lim_{t\to\infty}\lambda(t) = 0$.

Conditions (c) and (d) together imply

$$\lambda = \frac{e^{-(\rho+r)t}}{\rho+r},$$

which may be substituted in the conditions formulated in (a).

The Hamiltonian and necessary conditions for problem (10.2) look similar and therefore will not be written down explicitly.

Observe that, in general, interior solutions cannot always be expected. For an interior solution for α (and fixed β), the appropriate condition under (a) becomes

$$\frac{\partial g(\alpha,\beta)}{\partial \alpha} = 1 - \frac{e^{-rt}}{\rho+r},$$

and, assuming that the partial derivative of g with respect to α is nonnegative, this condition cannot be met for low values of t if $\rho + r < 1$. In that case $\alpha = 1$ for low values of t. On the other hand, if $\rho + r \geq 1$, a necessary condition to have an interior solution α for all values of t (and β) is that $r > 0$ and the derivative $\partial g(\alpha,\beta)/\partial\alpha$ takes all values between $1-1/(\rho + r)$ and 1.

In the following subsections we consider a few specifications of g and h which enable us to derive exact solutions.

3.2 Bang-bang Solutions

The specification considered here allows 'bang-bang' solutions, that is solutions taking only the values 0 and 1, among the open-loop Nash equilibria. Let

$$g(\alpha,\beta) = \alpha(1 - \beta), \; h(\alpha,\beta) = \beta(1 - \alpha).$$

The conditions in (a) – (d) of the previous section lead to

$$\alpha(t) = 0 \;\; \textit{if} \;\; \beta(t) > \frac{e^{-rt}}{\rho+r},$$

$$\alpha(t) = 1 \;\; \textit{if} \;\; \beta(t) < \frac{e^{-rt}}{\rho+r}$$

and analogous conditions for $\beta(t)$, depending on $\alpha(t)$. Further, α (or β) may take on arbitrary values between 0 and 1 if we have an equality sign in any of these conditions. This leads to the following description of open-loop Nash equilibria. Here, t' is the value of t for which $e^{-rt}/(\rho + r) = 1$ and t'' is the value of t for which $e^{-st}/(\rho + s) = 1$. Observe that $t' < t''$ if $r > s > 0$, provided that t' and t'' exist.

Proposition 1 With the specifications $g(\alpha,\beta) = \alpha(1 - \beta)$ and $h(\alpha,\beta) = \beta(1 - \alpha)$ and for $r > s > 0$, the open-loop Nash equilibria are combinations of strategies α and β containing only isolated discontinuities and satisfying:

(i) For every $0 \leq t < t'$: $\alpha(t) = \beta(t) = 1$.
(ii) For $t = t'$: $\beta(t) = 1$ and $\alpha(t)$ is arbitrary.
(iii) For $t' < t < t''$: $\beta(t) = 1$, $\alpha(t) = 0$.
(iv) For $t \geq t''$: $\beta(t) = 1$ and $\alpha(t) = 0$ or $\beta(t) = 0$ and $\alpha(t) = 1$ or $\beta(t) = e^{-rt}/(\rho + r)$ and $\alpha(t) = e^{-st}/(\rho + s)$.

Thus, in this specification there are solutions taking on only the values 0 and 1. Solutions of this kind are usually called *bang-bang solutions*. Both players might start off (depending on the values of r and s relative to the common discount factor ρ) with full investment in the first (competitive) project. In the longer run, however, in equilibrium either one of the players invests fully in this project and the other one invests nothing or the investments of both players are between 0 and 1 but in the long run converge to 0. It should be noted that the solutions in Proposition 1 are formulated at each point t in time separately, so that the resulting strategies may be highly discontinuous. The first player to jump (necessarily) to zero investment is the one with the higher of the two rates r and s (as can be easily seen); at that point, it is no longer advantageous to compensate for the comparative 'unlearning' effect of investment in the first project (given that the other player still invests fully) by also investing in that project. Thus, the player with the higher of the two rates r and s is the first one 'to give up'. A plausible equilibrium would be one where after this event this player stays at a zero investment level, while his opponent stays at investment level 1.

A proof of Proposition 1 can be based on the necessary conditions stated in the previous section and will not be elaborated.

3.3 A Cobb–Douglas Case

In this subsection we assume specifications which also allow interior solutions of the players' maximisation problems, that is, open-loop Nash equilibria with investments which may be strictly between 0 and 1. To be precise, we take

$$g(\alpha,\beta) = \sqrt{\alpha(1 - \beta)}\,, \quad h(\alpha,\beta) = \sqrt{\beta(1 - \alpha)}\,.$$

The analysis of the general Cobb–Douglas case is more tedious but will not exhibit essentially different features.

The next proposition describes the open-loop Nash equilibria for the situation analogous to the one in Proposition 1. In order to make the description easier to digest, we first introduce some notation.

Assume $r \geq s > 0$ and $\rho + r \leq 1$. Then let

$$t_1 = -\frac{\ln(\rho + r)}{r} \quad t_2 = -\frac{\ln(\rho + s)}{s}$$

$$t_3 = -\frac{\ln\frac{1}{2}(\rho + r)}{r} \quad t_4 = -\frac{\ln\frac{1}{2}(\rho + s)}{s}.$$

It can be verified that $0 \leq t_1 \leq t_2 \leq t_4$ and that $t_1 \leq t_3 \leq t_4$.

For $t \geq 0$ define

$$v(t) = 1 - \frac{e^{-rt}}{\rho + r} \quad w(t) = 1 - \frac{e^{-st}}{\rho + s}$$

$$\alpha^*(t) = \frac{1 - 4w^2(t)}{1 - 16v^2(t)w^2(t)} \quad \beta^*(t) = \frac{1 - 4v^2(t)}{1 - 16v^2(t)w^2(t)}.$$

We can now state our proposition.

Proposition 2 Assume $r \geq s > 0$ and $\rho + r \leq 1$, and $g(\alpha,\beta) = \sqrt{\alpha(1 - \beta)}$, $h(\alpha,\beta) = \sqrt{\alpha(1 - \alpha)}$. The open-loop Nash equilibria are combinations of strategies α and β containing only isolated discontinuities and satisfying:

(i) For $0 \leq t < t_1$: $\alpha(t) = \beta(t) = 1$.
(ii) For $t = t_1$: $\beta(t) = 1$, $\alpha(t)$ arbitrary.
(iii) For $t_1 < t < t_2$: $\beta(t) = 1$, $\alpha(t) = 0$.
(iv) For $t_2 \leq t < t_4$, there are two cases. (a) If $t_3 \leq t_2$, then $\beta(t) = 1$, $\alpha(t) = 0$. (b) If $t_2 < t_3 < t_4$, then for $t_2 \leq t \leq t_3$ there are three possibilities: $\beta(t) = 1$, $\alpha(t) = 0$, or $\beta(t) = 0$, $\alpha(t) = 1$ or $\beta(t) = \beta^*(t)$, $\alpha(t) = \alpha^*(t)$, while for $t_3 < t < t_4$: $\beta(t) = 1$, $\alpha(t) = 0$.
(v) For $t \geq t_4$: $\beta(t) = \beta^*(t)$, $\alpha(t) = \alpha^*(t)$.

Proposition 2 describes the most general case: in all other cases with $r \geq s$, the only difference may be that the whole picture moves to the left (or, equivalently, the origin to the right). Of course, the analysis of the case $r \leq s$ is similar. The proof of Proposition 2 is again based on the conditions formulated in the previous section, and will not be given in detail.

Figure 10.1 Proposition 2

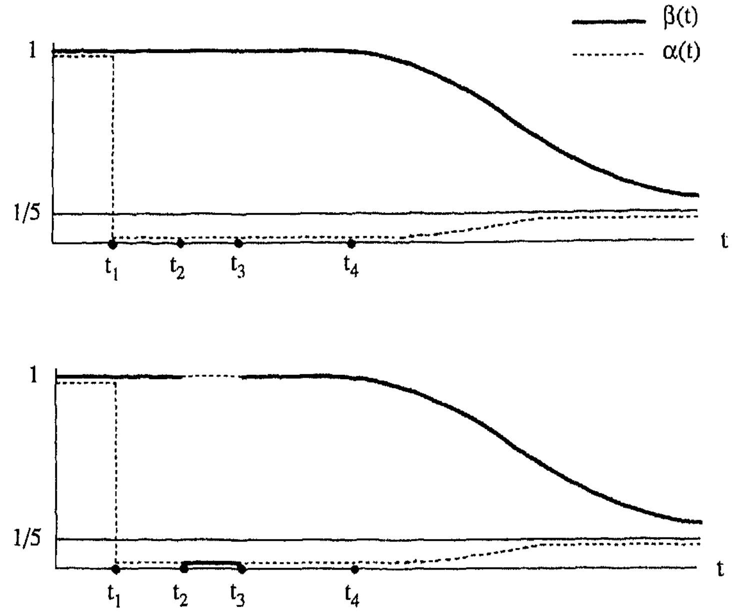

Figure 10.1 depicts some strategy combinations described by Proposition 2. Again, it should be noted that the strategies are defined for each t separately, and thus may contain any number of isolated discontinuities. In all cases, after t_4 the additional payoff effects associated with the state variables x and y decrease rapidly, causing player 1 to start investing again, while player 2 gradually decreases investments in the competitive project. In the limit, both α^* and β^* approach 1/5.

The following proposition applies to the situation where one of the two players has a nonpositive depreciation rate of the (un)learning payoffs.

Proposition 3 Assume $g(\alpha,\beta) = \sqrt{\alpha(1-\beta)}$, $h(\alpha,\beta) = \sqrt{\beta(1-\alpha)}$, and $r \leq 0$, $s > 0$, $\rho + r \leq 1$, and $|r| < \rho$. Let, as above, $t_2 = -\ln(\rho + s)/s$. Then, for an open-loop Nash equilibrium we have:

(i) For all $t > \max\{0,t_2\}$: $\alpha(t) = 1$, $\beta(t) = 0$.
(ii) For all $0 \leq t < t_2$: $\alpha(t) = \beta(t) = 1$.

This proposition confirms the obvious intuition that the player with the nonpositive depreciation rate survives, as far as investment in the first project is concerned. If $\rho + r > 1$, then in the longer run this will still hold although, initially, the equilibrium strategies may look different (details are omitted).

3.4 An Alternative Formulation

A drawback in the formulation of the investment problem in subsection 3.1 is that the effects of the state variables x and y vanish in the long run due to the presence of the coefficients e^{-rt} and e^{-st}; there is not only relative but also absolute 'unlearning' or 'depreciation' as time goes on.[3] To avoid this, we could alternatively require

$$x(t) = \int_0^t k(t,\tau) f(\alpha(\tau), \beta(\tau)) d\tau, \tag{10.3}$$

where, as before, x is the state variable for player 1, where f describes how the state variable depends on the investment plans of both players, and where the function k reflects the depreciation or growth of the state variable. Differentiating, we obtain

$$\dot{x}(t) = \int_0^t \frac{\partial k(t,\tau)}{\partial t} f(\alpha(\tau), \beta(\tau)) d\tau + k(t,t) f(\alpha(t), \beta(t)). \tag{10.4}$$

In the special situation that $\partial k(t,\tau)/\partial t = l(t)k(t,\tau)$ for some function l depending only on t, equation (10.4) implies

$$\dot{x}(t) = x(t)l(t) + k(t,t)f(\alpha(t), \beta(t)). \tag{10.5}$$

Instead of (10.1) now consider the maximisation problem

$$\begin{aligned} &\text{Maximise} \quad \int_0^\infty e^{-\rho t}[g(\alpha(t), \beta(t)) + (1-\alpha(t)) + x(t)]\,dt \\ &\text{subject to} \quad \dot{x}(t) = x(t)l(t) + k(t,t)f(\alpha(t),\beta(t)) \\ &\qquad\qquad\qquad x(0) = x_0, \alpha(t) \in [0,1]. \end{aligned} \tag{10.6}$$

The coefficient e^{-rt} has now been removed from the objective function; instead the state equation has been replaced by (10.5). A similar formulation can be given for player 2. The corresponding Hamiltonian is now given by

$$\begin{aligned} H(\alpha,x,t,\lambda) &= e^{-\rho t}[g(\alpha(t), \beta(t)) + (1-\alpha(t)) + x(t)] \\ &\quad + \lambda(t)[x(t)l(t) + k(t,t)f(\alpha(t), \beta(t))], \end{aligned}$$

with

$$\dot{\lambda} = -\frac{\partial H}{\partial x} = -e^{-\rho t} - l(t)\lambda(t)$$

as the costate equation. Again, similar expressions hold for player 2.

A simple example is obtained by assuming, in line with the preceding subsections,

$$f(\alpha(\tau),\ \beta(\tau)) = \alpha(\tau) - \beta(\tau),$$

and

$$k(t,\tau) = e^{r(\tau - t)}.$$

Then $l(t) = -r$ (we assume $r > 0$), and it is easily established, also using the transversality condition $\lim_{t\to\infty}\lambda(t) = 0$, that

$$\lambda(t) = \frac{e^{-\rho t}}{r+\rho}.$$

Substituting this expression for $\lambda(t)$ in the Hamiltonian and maximising at a given t and for a given strategy $\beta(t)$ of player 2 over the possible values of $\alpha(t)$, it follows easily that the maximising value of α will be independent of t. In other words (and making similar assumptions about player 2), in an open-loop Nash equilibrium the strategies of the players can be chosen constant over time, for this particular choice of the function k (and the corresponding function for player 2). For particular choices of the functions g and h (the immediate payoff functions of players 1 and 2 respectively, from investing in the first project), such an open-loop Nash equilibrium can be calculated, for instance for Cobb–Douglas payoff functions as in the preceding sections. Details are left to the reader.

4. SOME CONCLUDING REMARKS

In the foregoing, some attempts were made to study (un)learning effects in continuous-time two-person games. Here, (un)learning was to be understood in a 'physical' sense of acquiring certain skills in actions, not in the sense of (un)learning how to play the game. The main model was simple enough to enable the derivation of explicit solutions. The problem is that only slightly more sophistication in the model is bound to lead to mathematical intractibility as far as finding explicit analytical solutions is concerned.

NOTES

1. See also Cheng (1984) on this topic.
2. In what follows it is convenient to suppress t from the notation whenever this does not lead to confusion.
3. This was also pointed out to us by Fernando Vega-Redondo of Alicante University.

REFERENCES

Acharya, S.N. (1981), 'Perspectives and Problems of Development in Sub-Saharan Africa', *World Development*, 9, 109–146.

Arrow, K.J. (1962), 'The Economic Implications of Learning by Doing', *Review of Economic Studies*, 29, 155–173.

Cheng, L. (1984), 'International Competition in R&D and Technological Leadership', *Journal of International Economics*, 17, 15–40.

Heyer, J. (1976), 'Achievements, Problems and Prospects in the Agricultural Sector', in J. Heyer, J.K. Maitha and W.M. Senga (eds), *Agricultural Development in Kenya*, Oxford: Oxford University Press.

Joosten, R., Peters H. and Thuijsman F. (1991), 'Unlearning by not-doing: repeated games with vanishing actions', MERIT Research Memorandum No. 91–024, Maastricht.

McMillan, J. (1986), *Game Theory in International Economics*, Chur: Harwood Academic Publishers.

Mertens, J.F. and Neyman, A. (1981), 'Stochastic Games', *International Journal of Game Theory*, 10, 53–66.

Shapley, L.S. (1953), 'Stochastic Games', *Proceedings of the National Academy of Sciences USA*, 39, 1095–1100.

Starr, A.W. and Ho, Y.C. (1969a), 'Nonzero-sum Differential Games', *Journal of Optimization Theory and Applications*, 3, 184–206.

Starr, A.W. and Ho, Y.C. (1969b), 'Further Properties of Nonzero-sum Differential Games', *Journal of Optimization Theory and Applications*, 3, 207–219.