Differential Games with Random Terminal Instants

Sergey Kostyunin

Faculty of Applied Mathematics, St. Petersburg State University,
University pr., 35, Petrodvorets, 198904, St.Petersburg, Russia
E-mail: kostyunin.sergey@gmail.com

Abstract We investigate a noncooperative differential game with two players. Each player has his own random terminal time. After the first player leaves the game, the remaining one continues and gets the final reward for winning. An example is introduced where two firms compete in extracting a unique nonrenewable resource over time. The optimal feedback strategy, i.e. the optimal extraction rate, is calculated in a closed form.

Keywords: Differential game, random terminal time, Hamilton-Jacobi-Bellman equation

1. Introduction

In the last decades many economic models have been investigated with the precious help of the tools provided by differential game theory (see Dockner et al., 2000, Jørgensen and Zaccour, 2007). Both deterministic and stochastic approaches have been widely developed in a wide range of different frameworks.

This paper aims to analyze a class of models of differential games with 2 players. In particular, we consider a framework where the terminal instants of the game are random variables having different cumulative distribution functions. The first player which stops the game is the loser, whereas the remaining player gets a terminal reward and keeps playing. In this case the game collapses into a optimal control problem.

We are going to fully characterize the structure of the game and to determine its dynamic equilibrium structure. Finally, we will feature an example which is a modification of the standard model of extraction (see Rubio, 2006), with linear state dynamics and a logarithmic payoff structure. It will be completely discussed and its optimal feedback solution will be exhibited.

2. Game Formulation

There are two players which participate in differential game $\Gamma(t_0, x_0)$. The game $\Gamma(x_0)$ with dynamics

$$\dot{x} = g(t, x, u_1, u_2), \quad x \in R^n,\, u_i \in U \subseteq \operatorname{comp} R^l, \tag{1}$$
$$x(t_0) = x_0$$

starts from initial state x_0 at the time instant t_0. But here we suppose that each player has a distinct terminal time. The payoff of the game is composed of two components: the integral payoff achieved while playing, and the final reward, assigned to the player which stays alive after the retirement of its rival;

Let T_1 and T_2 be the independent random variables denoting the respective terminal instants of the players, and assume that their c.d.f. $F_1(\cdot)$, $F_2(\cdot)$ and their p.d.f. $f_1(\cdot)$ and $f_2(\cdot)$ are known. Random variables aren't bounded from above, i.e. $T_k \in [t_0; +\infty)$,
$k = 1, 2$.

Suppose, that for all feasible controls of players, participating the game, there exists a continuous at least piecewise differentiable and extensible on $[t_0, \infty)$ solution of a Cauchy problem (1).

Denote the instantaneous payoff of player i at the time τ, $\tau \in [t_0, \infty)$ by $h_i(\tau, x(\tau), u_1, u_2)$, or briefly $h_i(\tau)$. Suppose, that for all feasible controls of players which participate the game, the instantaneous payoff function of each player is bounded, piecewise continuous function of time τ (piecewise continuity is treated as following: function $h_i(\tau)$ could have only finitely many point of discontinuity on each interval $[t_0, t]$ and bounded on this interval).

Thereby, the function $h_i(\tau)$ is Riemann integrable on every interval $[t_0, t]$, in other words for every $t \in [t_0, \infty)$ there exists an integral $\int\limits_{t_0}^{t} h_i(\tau)d\tau$.

So, we have that the expected integral payoff of the player i can be represented as the following mathematical expectation:

$$I_i(t_0, x_0, u_1, u_2) = \mathbb{E}\left[\int_{t_0}^{\min\{T_1, T_2\}} h_i(t)dt\right], \tag{2}$$

where $\mathbb{E}[\cdot]$ is the mathematical expectation of a function of a random vector (T_1, T_2).

Moreover, we suppose that at the final (random) moment of the game, if player i is the only one remaining in the game, he receives the terminal payoff $\Phi_i(x(T))$, where $\Phi_i(x(T))$ are continuous functions on R^m. Then the expected terminal payoff of the player i can be evaluated as:

$$S_i(t_0, x_0, u_1, u_2) = \mathbb{E}\left[\Phi_i(x(T_j))\mathbb{I}_{[T_i > T_j]}\right], \tag{3}$$

where $\mathbb{I}_{[\cdot]}$ is the indicator function and $\mathbb{E}[\cdot]$ is the mathematical expectation of a function of a random vector (T_1, T_2).

Then the total expected payoff of the player i is:

$$K_i(t_0, x_0, u_1, u_2) = \mathbb{E}\left[\int_{t_0}^{\min\{T_1, T_2\}} h_i(t)dt + \Phi_i(x(T))\mathbb{I}_{[T_i > T_j]}\right]. \tag{4}$$

3. Transformation of expected payoff

The total expected payoff (4) is difficult to use in order to find solutions of the game. The standard methods of solution, such as Pontryagin's maximum principle or finding the solution of Hamilton-Jacobi-Bellman equatoin, can not be applied. We need to transform the payoff (4) into the standard integral functional for infinite-horizon differential games.

3.1. Expected integral payoff

At first consider the expected integral payoff (2). We could rewrite it in the following form by the definition of mathematical expectation

$$I_i(t_0,x_0,u_1,u_2) = \int\int_{t_0}^{\min\{\tau_1,\tau_2\}} h_i(t)dt\, dF_{T_1,T_2}(\tau_1,\tau_2), \tag{5}$$

where $F_{T_1,T_2}(\tau_1,\tau_2)$ is the cumulative distribution function of the random vector (T_1,T_2).

Consider the following function of the random vector:

$$T = \min\{T_1,T_2\}.$$

Since the function min{} is a measurable function, then T is a random variable (Borovkov, 1999). Denote by $F(t)$ the cumulative distribution function of the random variable T. Using the cumulative distribution functions of the random variables T_1, T_2, we can write the expression for $F(t)$ in an explicit form (Kostyunin et al., 2011)

$$F(t) = 1 - (1 - F_1(t))(1 - F_2(t))$$

Mathematical expectation (5) could be represented in the equivalent form (Borovkov, 1999)

$$I_i(t_0,x_0,u_1,u_2) = \int\int_{t_0}^{\tau} h_i(t)dt\, dF(\tau).$$

Thus, we could consider the expected integral payoff (5) as the mathematical expectation of a function of a random variable T:

$$I_i(t_0,x_0,u_1,u_2) = \mathbb{E}\left[\int_{t_0}^{T} h_i(t)dt\right], \tag{6}$$

where $\mathbb{E}[\cdot]$ is the mathematical expectation of a function of a random variable T.

If the instantaneous payoff function is nonnegative$h_i(\tau,x,u_1,u_2), \forall\tau,x,u_1,u_2$ then the following equality holds (Kostyunin and Shevkoplyas, 2011)

$$I_i(t_0,x_0,u_1,u_2) = \int\limits_{t_0}^{\infty} h_i(\tau)(1 - F(\tau))d\tau. \tag{7}$$

If the instantaneous payoff function does not satisfy the condition of nonnegativity, (7) holds if the following condition is satisfied (Kostyunin and Shevkoplyas, 2011)

$$\lim_{T\to\infty}(F(T) - 1)\int\limits_{t_0}^{T} h_i(t)dt = 0. \tag{8}$$

Note, that for a nonnegative instantaneous payoff function $h_i(t)$ the existence of the integral in the right-hand side of (7) implies that (8) holds.

3.2. Expected terminal payoff

Consider the expected terminal payoff (3)

$$S_i(t_0, x_0, u_1, u_2) = \mathbb{E}\left[\Phi_i(x(T_j))\mathbb{I}_{[T_i > T_j]}\right],$$

The expectation in (3) could be expressed as the following Lebesgue-Stieltjes integral:

$$\mathbb{E}\left[\Phi_i(x(T_j))\mathbb{I}_{[T_i > T_j]}\right] = \int \Phi_i(x(t_j))\mathbb{I}_{[t_i > t_j]} dF_{T_1,T_2}(t_1, t_2). \tag{9}$$

Suppose that the function $\Phi_i(x)$ satisfies the condition of nonnegativity. In this case we can use the following theorem on iterated integrals (Borovkov, 1999)

Theorem 1 (Theorem on iterated integrals). *For a Borel function $g(x, y) \geq 0$, and independent random variables ξ_1 è ξ_2:*

$$\int g(x_1, x_2) dF_{\xi_1 \xi_2}(x_1, x_2) = \int \left[\int g(x_1, x_2) dF_{\xi_2}(x_2)\right] dF_{\xi_1}(x_1).$$

Using this theorem, we obtain the following expression for (9)

$$\int_{t_0}^{+\infty} \left[\int_{t_0}^{+\infty} \Phi_i(x(t_j))\mathbb{I}_{[t_i > t_j]} dF_i(t_i)\right] dF_j(t_j).$$

Then we obtain

$$\int_{t_0}^{+\infty} \left[\int_{t_0}^{t_j} \Phi_i(x(t_j))\mathbb{I}_{[t_i > t_j]} dF_i(t_i) + \int_{t_j}^{+\infty} \Phi_i(x(t_j))\mathbb{I}_{[t_i > t_j]} dF_i(t_i)\right] dF_j(t_j).$$

The first term under the integral equals to zero. Further, we find

$$\int_{t_0}^{+\infty} \left[\int_{t_j}^{+\infty} \Phi_i(x(t_j)) dF_i(t_i)\right] dF_j(t_j) =$$
$$\int_{t_0}^{+\infty} \left[\Phi_i(x(t_j)) \int_{t_j}^{+\infty} f_i(t_i) dt_i\right] f_j(t_j) dt_j.$$

Finally, we obtain an expression for the expectation in (3)

$$\mathbb{E}\left[\Phi_i(x(T_j))\mathbb{I}_{[T_i > T_j]}\right] = \int_{t_0}^{+\infty} \Phi_i(x(t_j))(1 - F_i(t_j)) f_j(t_j) dt_j. \tag{10}$$

Then, the sufficient condition for total payoff transformation is given by the following propositions.

Proposition 1. *If the instantaneous payoff function and the terminal payment function are nonnegative*

$$h_i(\tau, x(\tau), u_1, u_2) \geq 0, \; \Phi_i(x(t)) \geq 0,$$

then the total expected payoff of player i (4) could be written as

$$K_i(t_0, x_0, u_1, u_2) = \int_{t_0}^{\infty} \left[h_i(\tau)\left(1 - F(\tau)\right) + \Phi_i(x(\tau)) f_j(\tau)(1 - F_i(\tau))\right] d\tau. \tag{11}$$

Proposition 2. *If the terminal payment function is nonnegative*

$$\Phi_i(x(t)) \geq 0,$$

and the following condition is satisfied

$$\lim_{T\to\infty} (F(T)-1) \int_{t_0}^{T} h_i(t)dt = 0,$$

then the total expected payoff of player i (4) could be written as (11).

4. Hamilton-Jacobi-Bellman equation

Let the game $\Gamma(t_0, x_0)$ develops along the trajectory $x(t)$. Then at the each time instant t, $t \in (t_0; \infty)$ players enter a new game (subgame) $\Gamma(t, x(t))$ with initial state $x(t) = x$.

The expected payoff for player i in this subgame is given by the following equation (Kostyunin et al., 2011)

$$K_i(t, x.u_1, u_2) =$$

$$\frac{1}{(1-F_1(t))(1-F_2(t))} \int_{t}^{+\infty} \left[h_i^*(\tau)\left(1-F(\tau)\right) + \Phi_i(x^*(\tau)) f_j(\tau)(1-F_i(\tau))\right] d\tau.$$

We denote by $W_i(t, x)$ the i-th optimal value function of the problem starting at $t \in (0, +\infty)$, with initial data $x(t) = x$. The Hamilton-Jacobi-Bellman equation has the same form as in the case where the terminal instants of the players are bounded from above (Kostyunin et al., 2011)

$$-\frac{\partial W_i(t,x)}{\partial t} + W_i(t,x)\left[\frac{f_1(t)}{1-F_1(t)} + \frac{f_2(t)}{1-F_2(t)}\right] =$$

$$\max_{u_i}[h_i(t, x, u_1, u_2) + \Phi_i(x(t))\frac{f_j(t)}{1-F_j(t)} + \frac{\partial W_i(t,x)}{\partial x}\phi(t, x, u_1, u_2)]. \qquad (12)$$

4.1. Hamilton-Jacobi-Bellman equation and hazard function

Let us remark that the term $\frac{f(\vartheta)}{1-F(\vartheta)}$ in the left-hand side of equation (12) is a well-known function in mathematical reliability theory. It has a name of Hazard function (or failure rate) with typical notation $\lambda(\vartheta)$

$$\lambda(t) = \frac{f(t)}{1-F(t)}. \qquad (13)$$

Using the definition of the Hazard function (13), we get the following form for new Hamilton-Jacobi-Bellman equation (12):

$$-\frac{\partial W_i(t,x)}{\partial t} + W_i(t,x)\left[\lambda_1(t) + \lambda_2(t)\right] =$$

$$\max_{u_i}[h_i(t, x, u_1, u_2) + \Phi_i(x(t))\lambda_j(t) + \frac{\partial W_i(t,x)}{\partial x}\phi(t, x, u_1, u_2)]. \qquad (14)$$

4.2. Exponential distribution case

For exponential distribution of terminal instants $F(t) = 1 - e^{-\lambda t}$, the Hazard function is constant: $\lambda(t) = \lambda$. So, inserting λ_i instead of $\lambda_i(t)$ into (12), we easily get the Hamilton-Jacobi-Bellman equation for player i

$$-\frac{\partial W_i(t,x)}{\partial t} + W_i(t,x)\left[\lambda_1 + \lambda_2\right] =$$
$$\max_{u_i}[h_i(t,x,u_1,u_2) + \Phi_i(x(t))\lambda_j + \frac{\partial W_i(t,x)}{\partial x}\phi(t,x,u_1,u_2)]. \qquad (15)$$

5. An example

Consider the following framework, borrowed from (Rubio, 2006) (Example 2.1) and (Dockner et al., 2000) (Example 5.7) and modified with the above discount factor. This example originally describes the joint exploitation of a pesticide, but its structure makes it suitable for our aim. Note that, in contrast to (Rubio, 2006), we confine our attention to the Nash equilibrium under simultaneous play, and we consider the non-stationary feedback case, that is our optimal value function explicitly depends on the initial instant t.

We fix $m = 1$, i.e., a unique state variable $x(t)$, denoting the amount of the resource, whereas the i-th payoff function explicitly depends on the rate of extraction of the i-th player but not on the state variable:

$$h_i(x(t), u_i(t)) = \ln u_i(t),$$

whereas the terminal payoff is given by

$$\Phi_i(x^*(T)) = c_i \ln(x(T_i)).$$

Note that $h_i(\cdot)$ is well-defined and concave for $u_i > 0$.

The transition function is linear and decreasing in the controls, so the dynamic constraint is:

$$\begin{cases} \dot{x} = -u_1 - u_2 \\ x(0) = x_0 > 0 \end{cases}.$$

The kinematic equation ensures that the terminal payoff is well-defined in that the resource cannot equal 0 in finite time.

Using the data of the above model, we obtain:

$$W_i(0,x_0) = \mathbb{E}\left[\int_0^{T_i} \ln u_i^* dt I_{[T_i<T_j]} + \int_0^{T_j} \ln u_i^* dt I_{[T_i>T_j]} + c_i \ln x(T_j)\mathbb{I}_{[T_i>T_j]}\right].$$

The i-th optimal value function of the problem starting at $t \in (0,\omega)$, and with initial condition $x(t) = x$, is given by:

$$W_i(t,x) =$$
$$\frac{1}{(1-F_i(t))(1-F_j(t))}\int_t^{\omega}\left[\ln u_i^*(\tau,x(\tau))\left(1-F(\tau)\right) + c_i \ln x(\tau) f_j(\tau)(1-F_i(\tau))\right] d\tau. \qquad (16)$$

In compliance with the previous Section, the Hamilton-Jacobi-Bellman equations are given by:

$$-\frac{\partial W_i(t,x)}{\partial t} + W_i(t,x)\left[\lambda_i(t)+\lambda_j(t)\right] = \max_{u_i}\left[\ln(u_i) + c_i \ln x(t)\lambda_j(t) - \frac{\partial W_i(t,x)}{\partial x}(u_i+u_j^*)\right]. \tag{17}$$

In order to explicitly determine the optimal strategy in the feedback Nash structure, we guess the following ansatz for the solution to (17):

$$W_i(t,x) = A_i(t)\ln x + B_i(t),$$

where $A_i(t)$ and $B_i(t)$ are unknown functions of t, such that the following limits are satisfied:

$$\lim_{t\to\omega} A_i(t) = 0, \qquad \lim_{t\to\omega} B_i(t) = 0. \tag{18}$$

The relevant first order partial derivatives to be employed in (17) are:

$$\frac{\partial W_i(t,x)}{\partial t} = \dot{A}_i(t)\ln x + \dot{B}_i(t), \qquad \frac{\partial W_i(t,x)}{\partial x} = \frac{A_i(t)}{x}.$$

Maximizing the r.h.s. of (17) yields:

$$\frac{1}{u_i^*} - \frac{\partial W_i(t,x)}{\partial x} = 0 \iff u_i^* = \frac{x}{A_i(t)}.$$

Hence, plugging u_i^*, $\frac{\partial W_i(t,x)}{\partial t}$ and $\frac{\partial W_i(t,x)}{\partial x}$ into (17), we obtain the following equation:

$$-\dot{A}_i(t)\ln x - \dot{B}_i(t) + (A_i(t)\ln x + B_i(t))\left[\lambda_i(t)+\lambda_j(t)\right] = \ln\frac{x}{A_i(t)} + c_i\ln x\lambda_j(t) - \frac{A_i(t)}{x}\left(\frac{x}{A_i(t)} + \frac{x}{A_j(t)}\right). \tag{19}$$

After collecting terms with and without $\ln x$, we determine the following ODEs for the time-dependent coefficients of $W_i(t,x)$:

$$-\dot{A}_i(t) + A_i(t)\left[\lambda_i(t)+\lambda_j(t)\right] - 1 - c_i\lambda_j(t) = 0, \tag{20}$$

$$-\dot{B}_i(t) + B_i(t)\left[\lambda_i(t)+\lambda_j(t)\right] + \ln A_i(t) + 1 + \frac{A_i(t)}{A_j(t)} = 0, \tag{21}$$

composing a Cauchy problem endowed with the transversality conditions:

$$\lim_{t\longrightarrow\omega} A_i(t) = 0, \qquad \lim_{t\longrightarrow\omega} B_i(t) = 0. \tag{22}$$

Proposition 3. *The optimal feedback strategy for the i-th firm is given by:*

$$u_i^*(t,x) = \frac{x}{\int_t^\omega (1+c_i\lambda_j(\tau))e^{-\int_t^\tau(\lambda_i(\theta)+\lambda_j(\theta))d\theta}d\tau}. \tag{23}$$

Proof. We just consider the Cauchy problem in $A_i(t)$, because the explicit calculation of $B_i(t)$ can be avoided in that $B_i(t)$ does not appear in the expression of u_i^*:

$$\begin{cases} \dot{A}_i(t) = A_i(t)\left[\lambda_i(t) + \lambda_j(t)\right] - 1 - c_i\lambda_j(t) \\ \lim_{t\to\omega} A_i(t) = 0 \end{cases},$$

whose general solution is given by:

$$A_i(t) = e^{\int_0^t (\lambda_i(\tau)+\lambda_j(\tau))d\tau}\left(C - \int_0^t (1 + c_i\lambda_j(\tau))e^{-\int_0^\tau (\lambda_i(s)+\lambda_j(s))ds}d\tau\right), \qquad (24)$$

where the constant C is determined by employing the transversality condition on $A_i(t)$:

$$C = \int_0^\omega (1 + c_i\lambda_j(\tau))e^{-\int_0^\tau (\lambda_i(s)+\lambda_j(s))ds}d\tau,$$

leading to the solution:

$$A_i^*(t) = e^{\int_0^t (\lambda_i(\tau)+\lambda_j(\tau))d\tau}\left[\int_t^\omega (1 + c_i\lambda_j(\tau))e^{-\int_0^\tau (\lambda_i(s)+\lambda_j(s))ds}d\tau\right]. \qquad (25)$$

We can simplify:

$$A_i^*(t) = \int_t^\omega (1 + c_i\lambda_j(\tau))e^{-\int_t^\tau (\lambda_i(s)+\lambda_j(s))ds}d\tau. \qquad (26)$$

Finally, the expression of the optimal feedback strategy for the i-th firm can be achieved from the FOCs of the model:

$$u_i^*(t,x) = \frac{x}{A_i^*(t)} = \frac{x}{\int_t^\omega (1 + c_i\lambda_j(\tau))e^{-\int_t^\tau (\lambda_i(\theta)+\lambda_j(\theta))d\theta}d\tau}. \qquad (27)$$

As a further application, we can consider the circumstance where the two distributions of the firms are the standard exponential distributions, i.e.

$$f_i(t;\ \lambda_i) = \begin{cases} \lambda_i e^{-\lambda_i t}, & \text{if } t \geq 0 \\ 0, & \text{if } t < 0 \end{cases},$$

whose means are respectively $\lambda_1^{-1},\ \lambda_2^{-1}$, both positive, with $\lambda_1 \neq \lambda_2$, ensuring asymmetry.

In this case the hazard functions are constant, i.e. $\lambda_1(t) \equiv \lambda_1$ and $\lambda_2 \equiv \lambda_2$, then substituting in (23) we obtain the two optimal feedback strategies:

$$u_1^*(t,x) = \frac{(\lambda_1 + \lambda_2)x}{(1 + c_1\lambda_2)[1 - e^{-(\lambda_1+\lambda_2)(\omega-t)}]}, \qquad (28)$$

$$u_2^*(t,x) = \frac{(\lambda_1 + \lambda_2)x}{(1 + c_2\lambda_1)[1 - e^{-(\lambda_1+\lambda_2)(\omega-t)}]}. \qquad (29)$$

6. Concluding remarks

This paper intends to be a contribution to the literature of differential games in an area which can be defined as deterministic, but enriched with some stochastic elements. In particular, it is focused on the feature of extraction games that is definitely realistic: the uncertainty about the terminal times of an extracting activity.

The dynamic feedback equilibrium structure has been determined and the specific technicalities of this setting have been pointed out. As an example, a model of nonrenewable resource extraction with a logarithmic utility structure was examined and solved in a closed form.

There exist some possible further extensions, also concerning the example we developed. It would be interesting to check the specific optimal strategies in presence of more complex hazard functions (for example, the Weibull distribution) or endowed with alternative payoff structures. Another interesting development might consist in considering a competition among more than 2 firms, having different terminal times.

References

Borovkov, A. A. (1999). *Probability Theory.* CRC Press, 484 p.

Dockner, E., S. Jørgensen, N. Van Long, G. Sorger (2000). *Differential games in economics and management science.* Cambridge University Press.

Jørgensen, S. and G. Zaccour (2007). *Developments in Differential Game Theory and Numerical Methods: Economic and Management Applications.* Computational Management Science, **4(2)**, 159–182.

Kostyunin, S., A. Palestini and E. Shevkoplyas (2011). *Differential game of resource extraction with random time horizon and different hazard functions.* Control Processes and Stability: proceedings of XLII international conference / Ed. by A.S. Eremin, N.V. Smirnov. Saint-Petersburg, Saint-Petersburg State University Publishing House, 571–576.

Kostyunin, S. and E. Shevkoplyas (2011). *On simplification of integral payoff in the differential games with random duration.* Vestnik St. Petersburg University. Ser. 10, Issue 4, 47–56.

Marin-Solano, J. and E. Shevkoplyas (2011). *Non-constant discounting in differential games with random time horizon.* Automatica, vol.48, 50, 2011. DOI: 10.1016/j.automatica.2011.09.010

Petrosjan, L. A. and G. Zaccour (2003). *Time-consistent Shapley Value Allocation of Pollution Cost Reduction.* Journal of Economic Dynamics and Control, **27**, 381–398.

Petrosjan L. A. and E. V. Shevkoplyas (2003). *Cooperative Solutions for Games with Random Duration.* Game Theory and Applications, Vol. IX. Nova Science Publishers, 125–139.

Rubio, S. (2006). *On Coincidence of Feedback Nash Equilibria and Stackelberg Equilibria in Economic Applications of Differential Games.* Journal of Optimization Theory and Applications, **128(1)**, 203–221.