Switching Lines for Optimal Control in Differential Game with Two Pursuers and One Evader

S.A. Ganebny, S.S. Kumkov, V.S. Patsko,
(Institute of Mathematics and Mechanics UrB RAS, Ekaterinburg, Russia)

S. Le Menec
(EADS / MBDA France, Paris, France)

The 15th International Symposium on
Dynamic Games and Applications
ISDG'2012

Liblice, Czech Republic, July 19 - 22, 2012

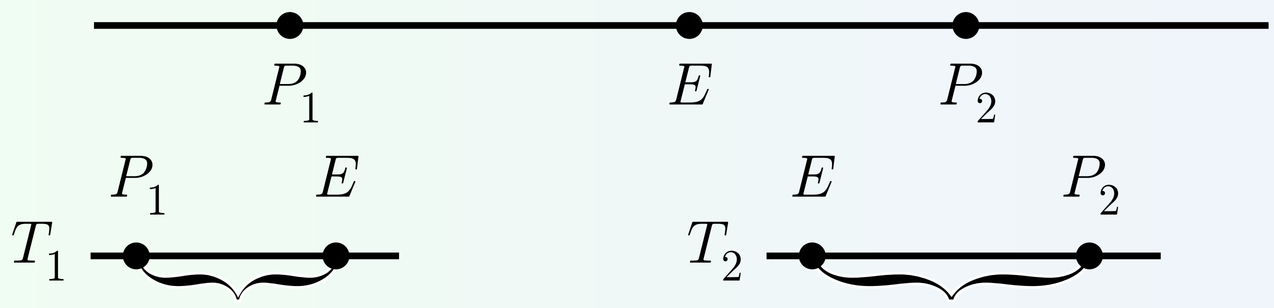
Model Pursuit Problem in a Straight Line

$$\ddot{z}_{P_i} = a_{P_i}, \qquad \ddot{z}_E = a_E,$$

$$\dot{a}_{P_i} = (u_i - a_{P_i})/l_{P_i}, \qquad \dot{a}_E = (v - a_E)/l_E,$$

$$|u_i| \le \mu_i,\ a_{P_i}(t_0) = 0,\ i = 1, 2, \qquad |v| \le \nu,\ a_E(t_0) = 0.$$

$$\varphi = \min\Big\{\big|z_E(T_1) - z_{P_1}(T_1)\big|, \big|z_E(T_2) - z_{P_2}(T_2)\big|\Big\}$$



Similar one-to-one problems and their practical interpretations have been considered by J.Shinar and his collaborators.

Le Menec S. Linear Differential Game with Two Pursuers and One Evader // Annals of the International Society of Dynamic Games, Vol.11, 2011, pp.209-226.

We have two pursuers and one evader. Geometric positions of all objects are on the straight line. On the left, one can see the dynamics of the first and second pursuers; on the right, there is the dynamics of the evader. We join the pursuers into the first player, the evader is the second player.

Fix two instants T_1 and T_2. At the instant T_1 (T_2), we compute the distance between the first (the second) pursuer and the evader. The minimum of these two distances is the resultant miss. The first player minimizes the miss. The interest of the second player is opposite. The first paper on this game was the work by S. Le Menec.

Description of Relative Dynamics

Change of variables: $y_1 = z_E - z_{P_1}, \qquad y_2 = z_E - z_{P_2}$

The dynamics is:

$$
\begin{aligned}
\ddot{y}_1 &= -a_{P_1} + a_E, & \qquad \ddot{y}_2 &= -a_{P_2} + a_E,\\
\dot{a}_{P_1} &= (u_1 - a_{P_1})/l_{P_1}, & \qquad \dot{a}_{P_2} &= (u_2 - a_{P_2})/l_{P_2},\\
\dot{a}_E &= (v - a_E)/l_E
\end{aligned}
$$

Here, y_1 and y_2 are the relative geometric coordinates; a_{P_1}, a_{P_2}, a_E are the accelerations; l_{P_1}, l_{P_2}, l_E are the time constants.

Constraints for the players' controls:

$$|u_1| \le \mu_1, \quad |u_2| \le \mu_2, \quad |v| \le \nu$$

The payoff function:

$$\varphi\big(y_1(T_1), y_2(T_2)\big) = \min\big(\big|y_1(T_1)\big|, \big|y_2(T_2)\big|\big)$$

Let us take geometric difference coordinates. As a result, we have antagonistic differential game, for which the dimension of the state vector is equal to 7.

Dynamics in Forecasted Coordinates (ZEM, *Zero Effort Miss* Coordinates)

Consider coordinates x_1 and x_2 that are the values of y_1 and y_2 forecasted to the respective termination instants T_1 and T_2 under zero controls:

$$x_i = y_i + \dot{y}_i \tau_i - a_{P_i} l_{P_i}^2 h(\tau_i / l_{P_i}) + a_E l_E^2 h(\tau_i / l_E),$$
$$\tau_i = T_i - t, \qquad h(\alpha) = e^{-\alpha} + \alpha - 1$$

The dynamics:

$$\dot{x}_1 = -l_{P_1} h(\tau_1 / l_{P_1}) u_1 + l_E h(\tau_1 / l_E) v,$$
$$\dot{x}_2 = -l_{P_2} h(\tau_2 / l_{P_2}) u_2 + l_E h(\tau_2 / l_E) v$$

The parameters of the game:

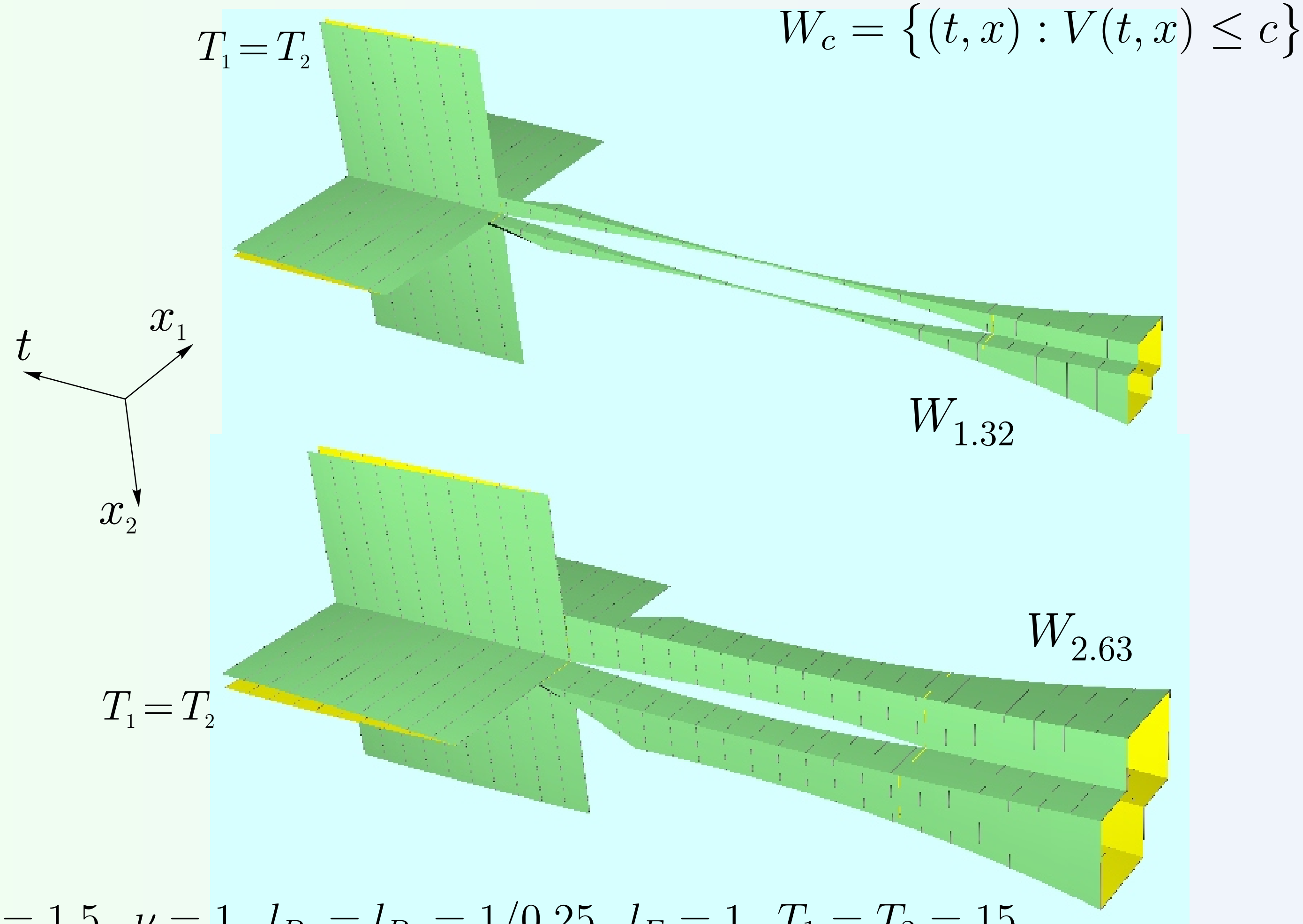
$$\eta_i = \frac{\mu_i}{\nu}, \qquad \varepsilon_i = \frac{l_E}{l_{P_i}}$$

η_i – maximal relative acceleration,

$\eta_i \varepsilon_i$ – agility

Because we fix termination instants T_1 and T_2, we can consider the prognosis miss between the first pursuer and the evader at the instant T_1, and between the second pursuer and evader at the instant T_2. Computing the prognosis miss, we suppose that the controls of the corresponding objects are equal to zero starting from the current time t. So, we use so-called "zero effort miss" scalar coordinates x_1 and x_2 for the equivalent differential game. There are no state variables in the right-hand side of the new dynamics. The dynamics has the parameters μ_1, μ_2, ν, l_{P1}, l_{P2}, and l_E. It is known that we can use normalized parameters η_1, η_2, ε_1, and ε_2.

Level Sets of the Value Function (Varying Advantage of Pursuers, Variant 1)



$\mu_1 = \mu_2 = 1.5,\ \nu = 1,\ l_{P_1} = l_{P_2} = 1/0.25,\ l_E = 1,\ T_1 = T_2 = 15.$

When we investigate some differential game, especially if we try to find optimal strategies of the players, then it is very useful to see the typical structure of level sets W_c of the value function V. In our game, we have three-dimensional sets in the space x_1, x_2, and t.

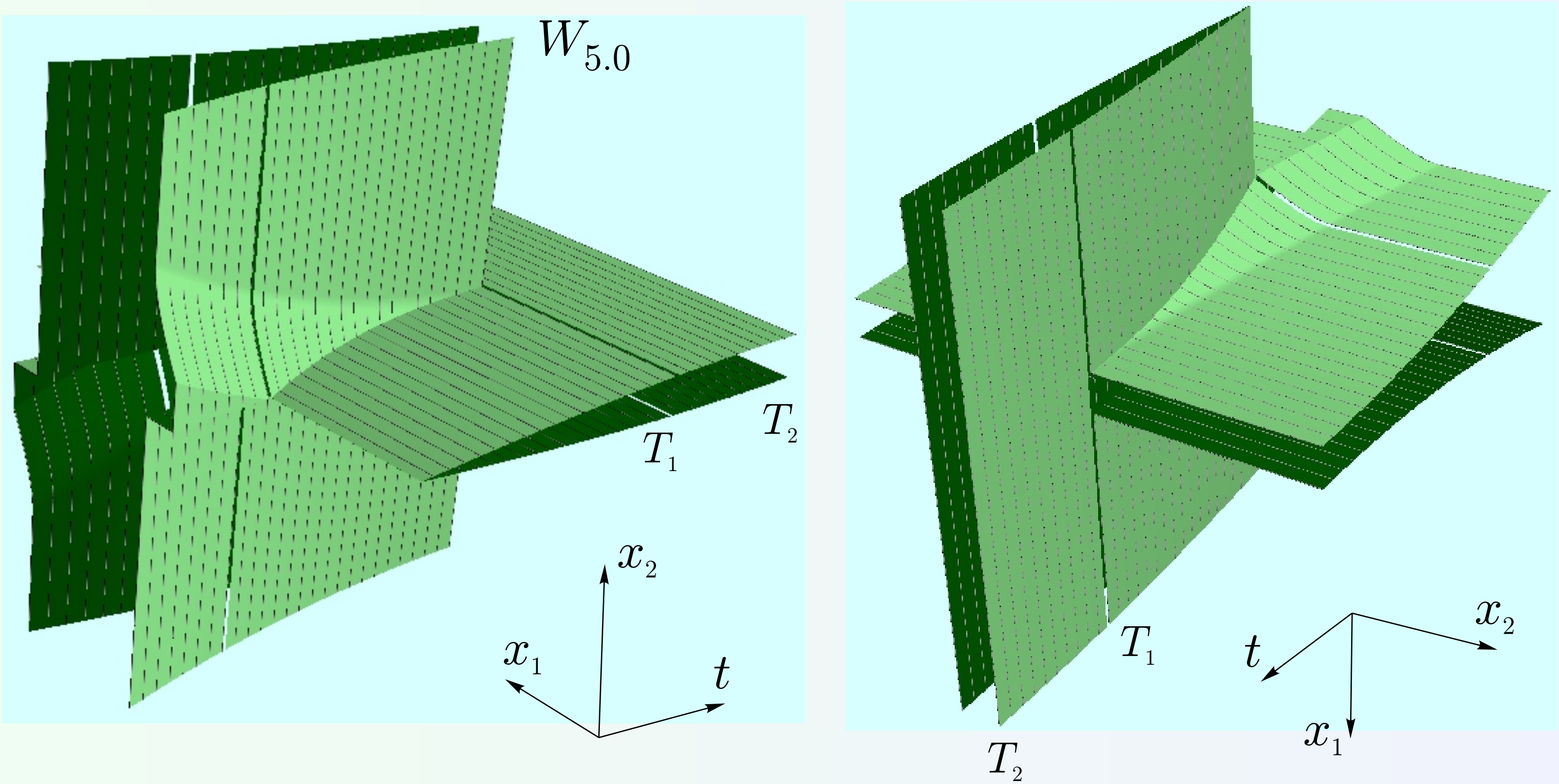
Here, we see two level sets for $c = 1.32$ and $c = 2.63$. These level sets were constructed numerically for the case of identical pursuers . When the remaining time-to-go is large, each of pursuers is stronger than the evader. If the time-to-go is small, the evader is more agile.

Theoretically, cross-like wings of the sets are infinite along the axes x_1 and x_2.

If the initial position of the game is inside of a level set, the first player guarantees guiding the system to the corresponding terminal cross. If the initial position is outside a tube, such a guarantee is absent. So, we can say that a level set of the value function is the solvability set for corresponding value of the payoff.

Level Set of the Value Function (One Strong and One Weak Pursuer)

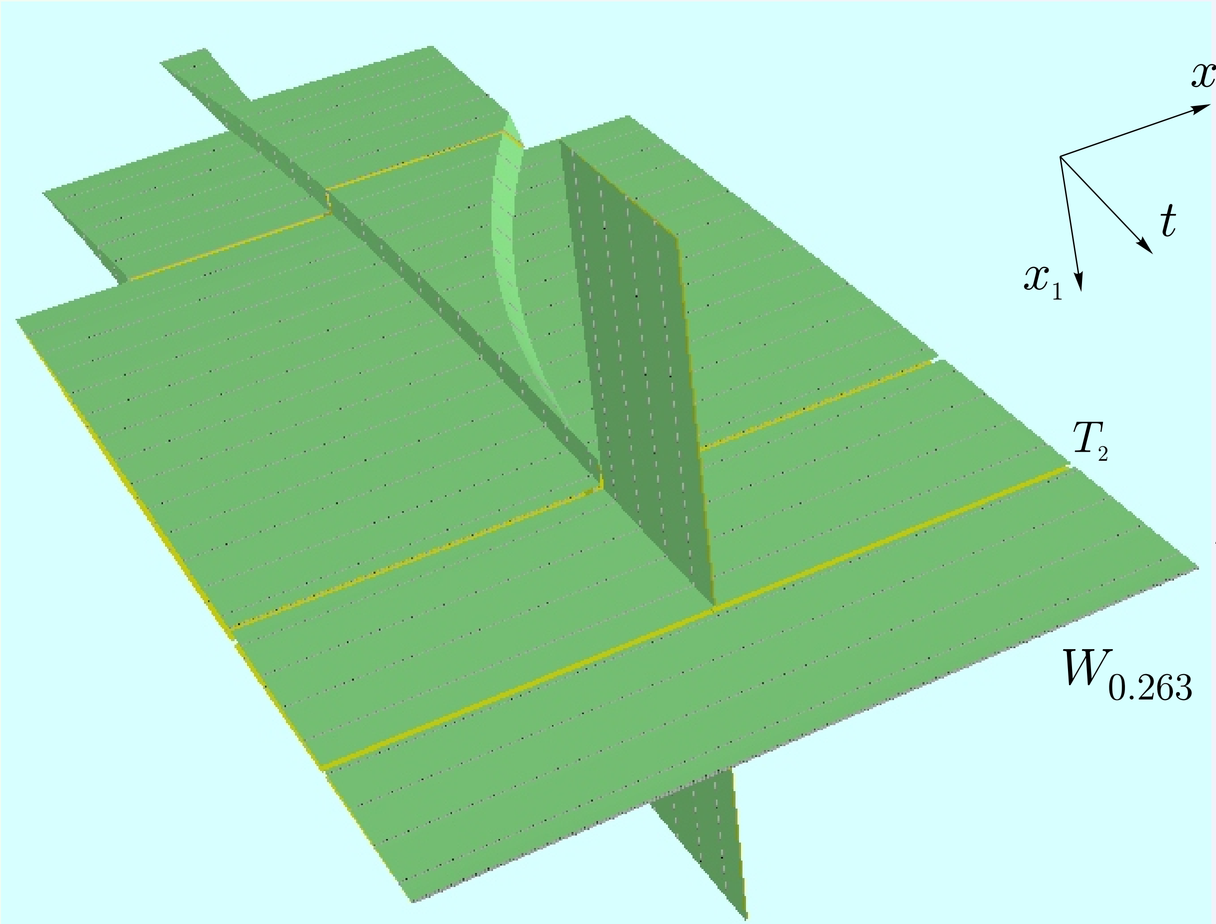
$$W_c = \{(t, x) : V(t, x) \le c\}$$



$$\mu_1 = 2.0,\ \ \mu_2 = 1.0,\ \ \nu = 1,\ \ l_{P_1} = 1/2.0,\ \ l_{P_2} = 1/0.3,\ \ l_E = 1,\ \ T_1 = 5,\ \ T_2 = 7.$$

Here, we see one level set of the value function, but from two points of view. This is for the case of one strong and one weak pursuers. Terminal instants T_1 and T_2 are not equal. These figures of level set and all others in our presentation are made with the help of special designer for 3D-pictures elaborated by Sergey Kumkov and his students. We clearly see where the value function is not differentiable.

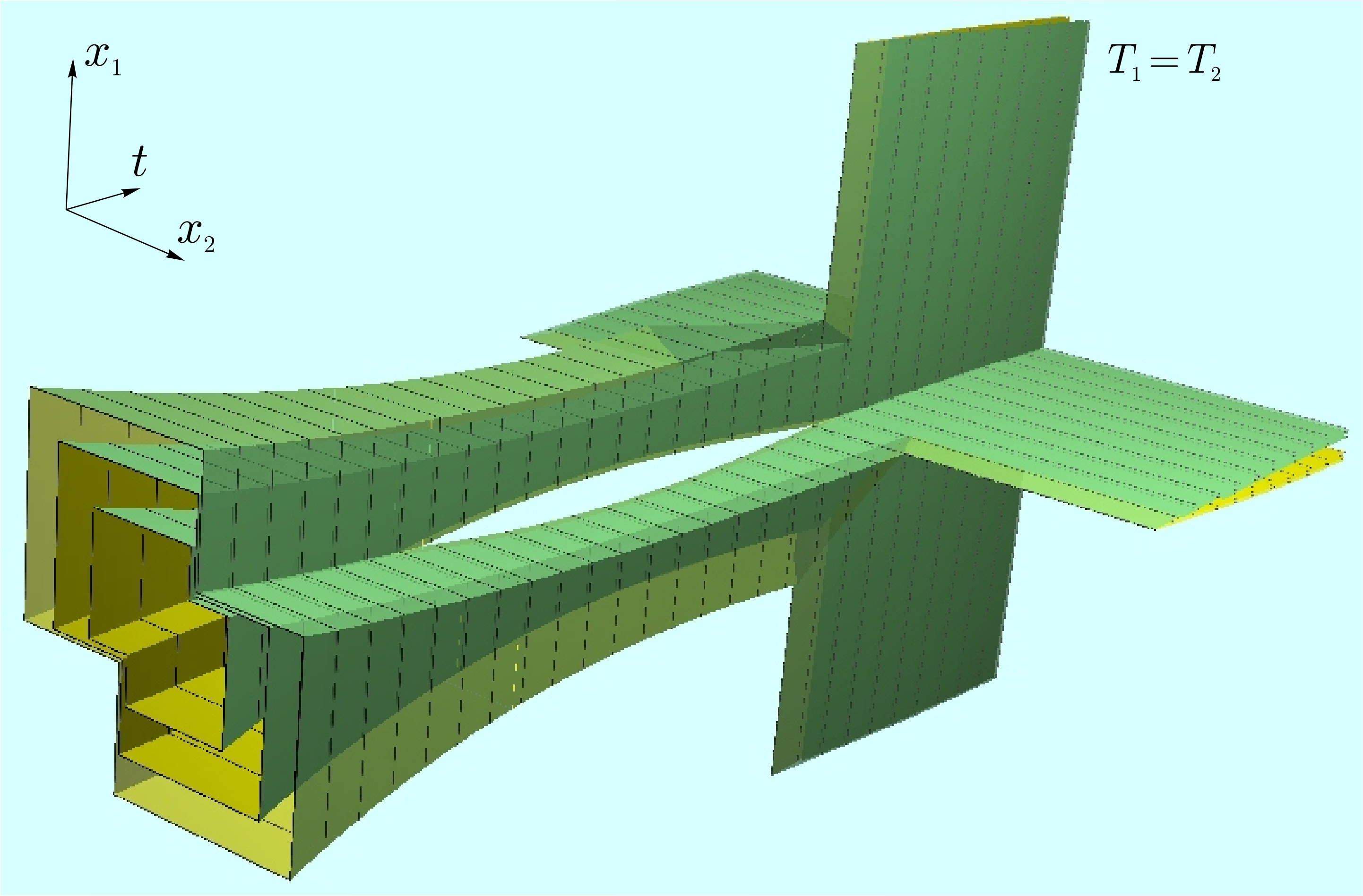
Level Set of the Value Function (Varying Advantage of Pursuers, Variant 2)



$$\mu_1 = 0.8,\ \ \mu_2 = 1.3,\ \ \nu = 1,\ \ l_{P_1} = 1/20,\ \ l_{P_2} = 1/0.5,\ \ l_E = 1,\ \ T_1 = 15,\ \ T_2 = 13.5.$$

In this slide, the set $W_{0.263}$ is shown for another case of varying advantage of pursuers. In the theory, the extension of the set along the axis x_1 (x_2) is infinite near the instants T_2 (T_1) of the direct time. For instants far from the terminal ones, the time sections of the level set are bounded.

Level Sets of the Value Function (Varying Advantage of Pursuers, Variant 1)

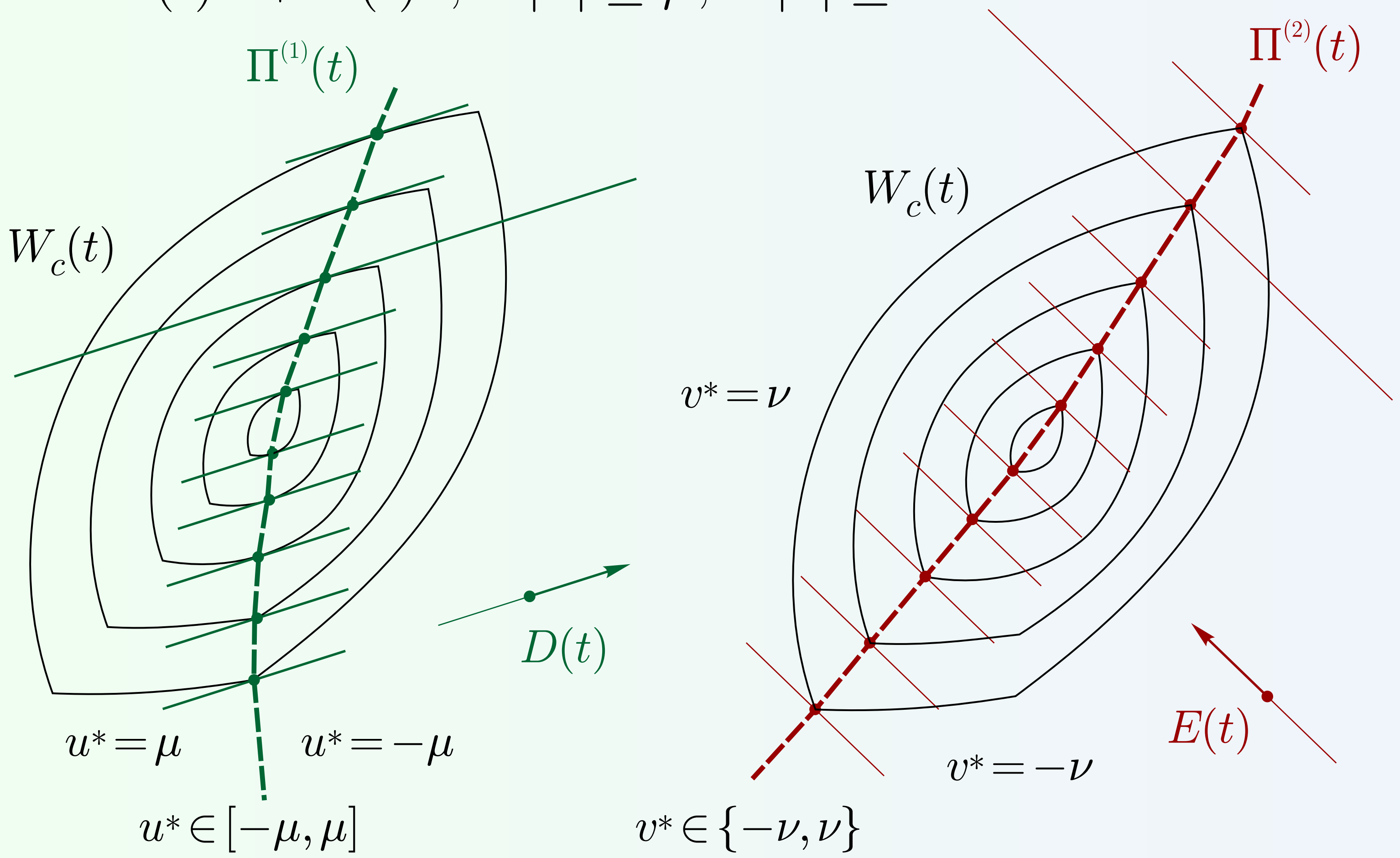


$\mu_1 = \mu_2 = 1.5,\ \nu = 1,\ l_{P_1} = l_{P_2} = 1/0.25,\ l_E = 1,\ T_1 = T_2 = 15.$

Of course, when we want to build optimal or quasioptimal feedback strategies of the players, we can use for this aim some collection of the level sets W_c of the value function computed numerically. But storing such three-dimensional sets is memory-consuming and inelegant.

Constructing Switching Lines

$$\dot{x} = D(t)u + E(t)v, \quad |u| \le \mu, \quad |v| \le \nu$$



N.D.Botkin, V.S.Patsko, M.A.Zarkh

What do we do in the case of linear game of the second order with fixed termination time and convex payoff function? Let the equivalent game have a form as above and controls u and v be scalar. So, $D(t)$ and $E(t)$ are vectors.

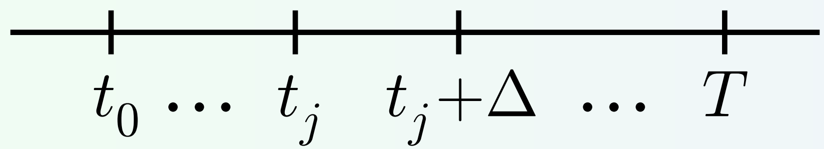
Processing t-sections of level sets W_c with a help of the vector $D(t)$ gives us the switching line $\Pi^{(1)}(t)$ of the first player. On one side from this line, the optimal control u^* is equal to $-\mu$, on the other side it equals $+\mu$. On the switching line, the optimal control can take arbitrary value from the interval $[-\mu, +\mu]$. For any straight line in the plane x_1, x_2 parallel to the vector $D(t)$, the switching point is the point of minimum of the restriction of the value function to this line.

Similar situation is for the second player. Here, we use the vector $E(t)$. But in contrast to the case of the first player, the optimal value v^* on the switching line can be only $-\nu$ or $+\nu$, not an intermediate value. For any straight line in the plane x_1, x_2 parallel to the vector $E(t)$, the switching point is the point of minimum of the restriction of the value function to this line too.

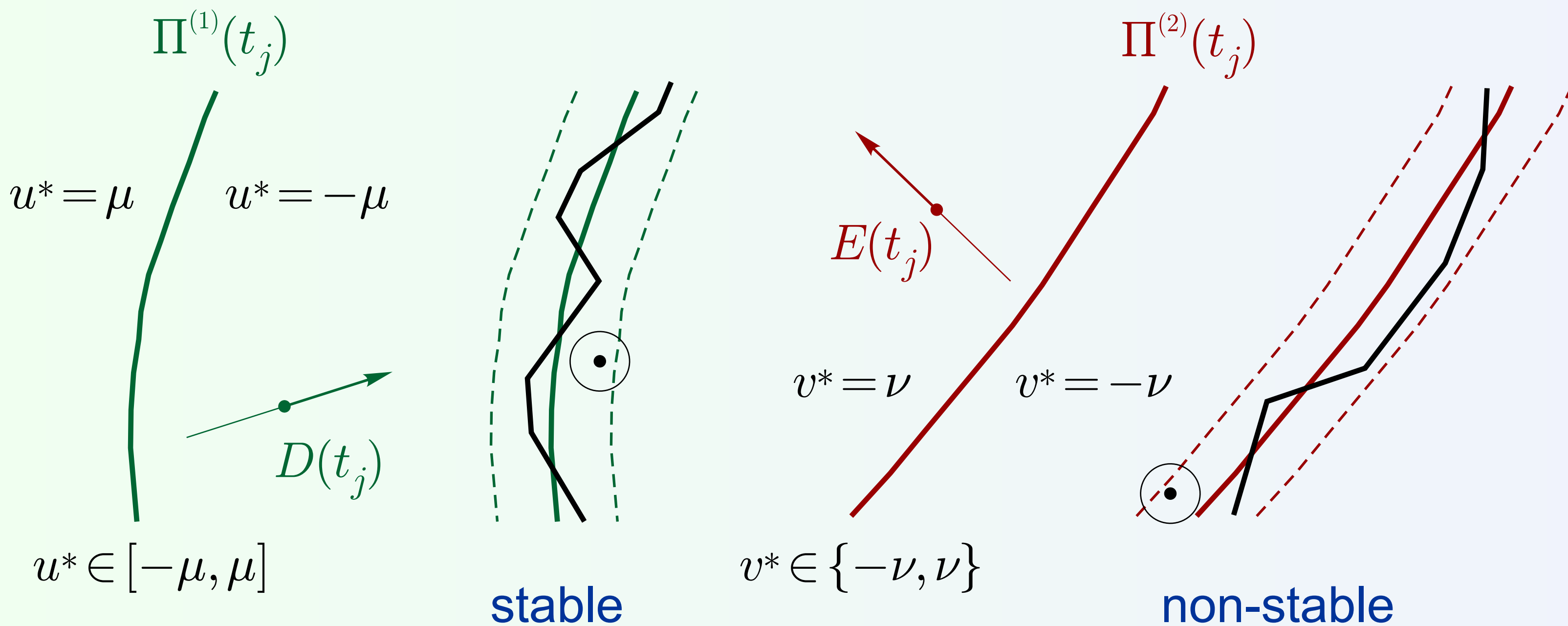
Problems concerning optimal feedback control on the basis of switching lines were investigated in detail by Nikolai Botkin, Valerii Patsko, and Michael Zarkh.

The main idea of the method is the following. We construct numerically level sets of the value function on some grid of values c, then we process them to obtain switching lines on some grid of instants t. We keep in memory the switching lines only and construct feedback controls of the first and second players using these lines.

Krasovskii's Discrete Control Scheme. Stability of Optimal Feedback Control



$u(t) = u^*\big(t_j, x(t_j)\big),\ t \in [t_j, t_j+\Delta)$ $\qquad$ $v(t) = v^*\big(t_j, x(t_j)\big),\ t \in [t_j, t_j+\Delta)$



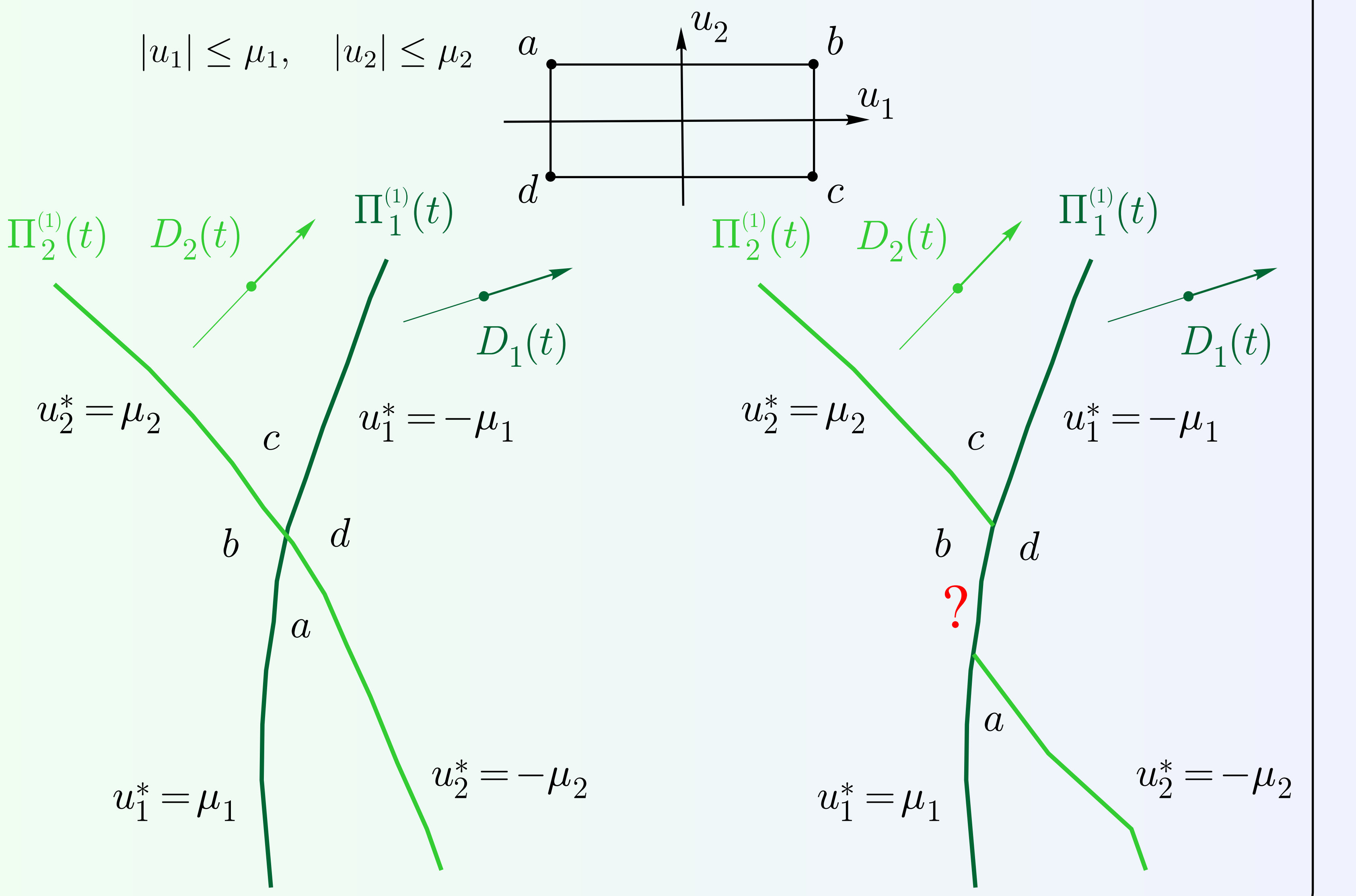
We use Krasovskii's formalization of differential games. In this formalization, a feedback control is used in discrete control scheme. Namely, at first, the entire time interval of the game is covered by a grid of instants. Then, being at some instant t_j, a player chooses its control and keeps it constant during the next time interval $[t_j, t_{j+1}) = [t_j, t_j+\Delta)$.

In the convex case, the situation for the first player is very good. Optimal feedback control is stable with respect to inaccuracies of numerical constructions of switching lines and with respect to information errors in measurements of state position of the system.

But this method is non-stable for the second player. However, non-stability can appear only under very skillful behavior of his opponent, that is, of the first player. If such a behavior is applied, the motion of the system goes in time along the switching surface of the second player with decreasing game value.

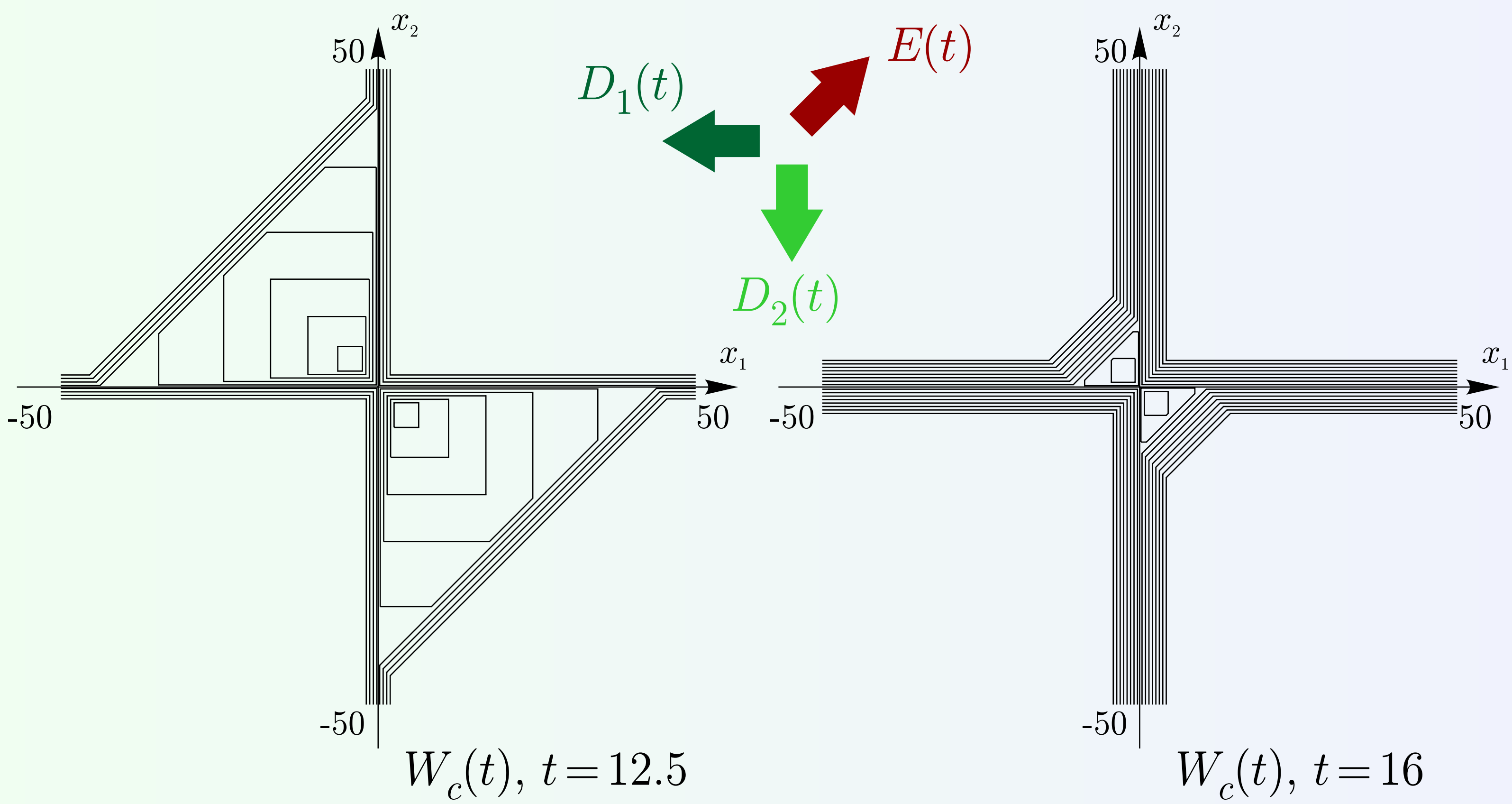
Case of Vector Control

$$|u_1| \le \mu_1, \quad |u_2| \le \mu_2$$



For the case of vector control with components bounded independently, we can construct a collection of switching lines, each one for a certain component of the control using the corresponding vector $D_i(t)$, which is the corresponding column of the matrix $D(t)$. But in this situation, a worse thing can occur, namely, confluence of two or more switching lines. In this slide, the control u of the first player is two-dimensional, constrained by a rectangle. $D_1(t)$ and $D_2(t)$ are the first and second columns of the matrix $D(t)$. The situation shown by the left figure is good. But the situation at the right is bad, because there is a coinciding part of lines $\Pi_1^{(1)}(t)$ and $\Pi_2^{(1)}(t)$. If a chattering regime of the first player occurs near this part with switching the control from b to d and back, the system motion can go near the switching surface with increasing of the game value. But in this case, the second player should behave itself very specifically.

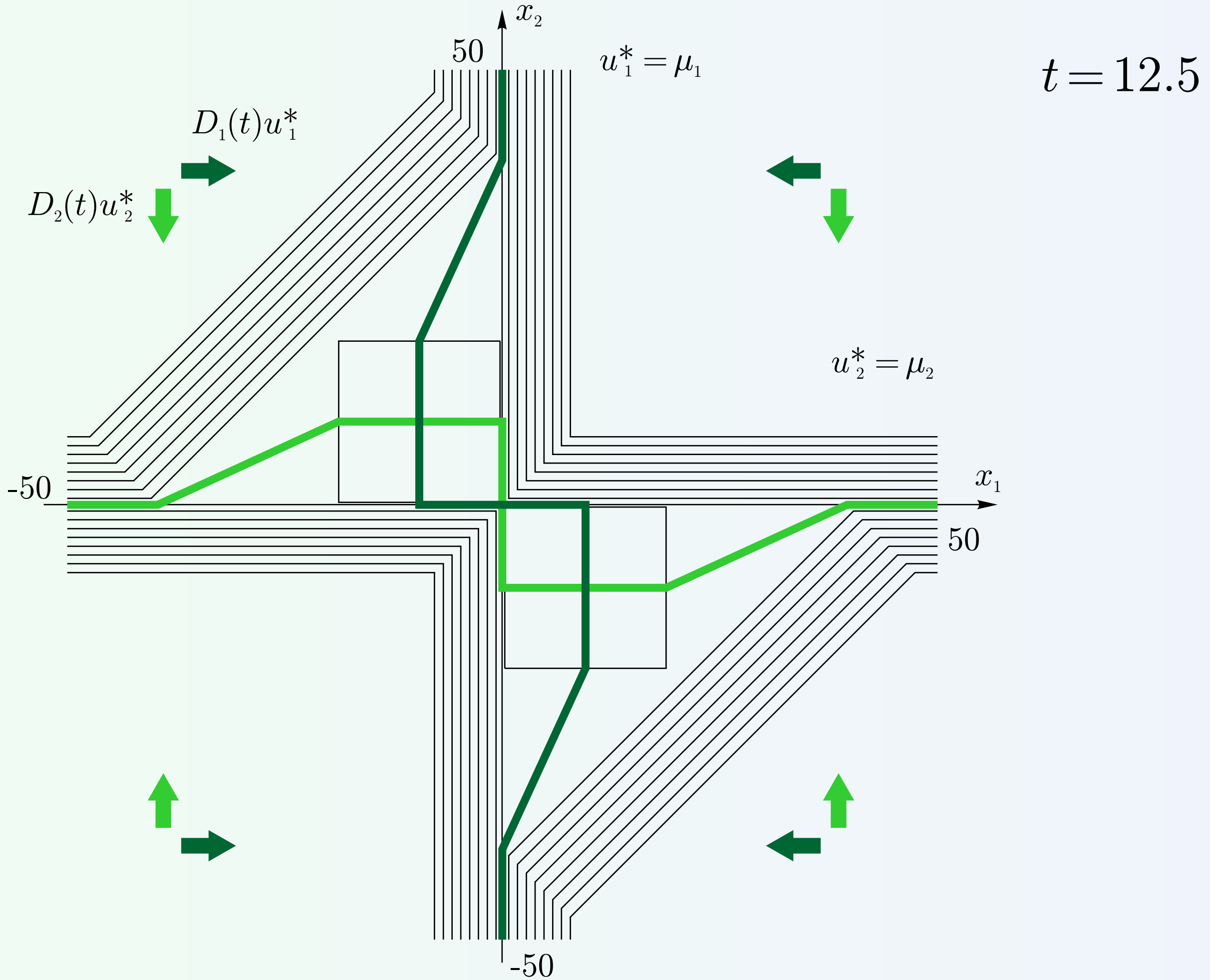
Time Sections of Level Sets of the Value Function (Varying Advantage of Pursuers, Variant 1)



In our problem with two pursuers and one evader, the payoff function is not convex. Due to this, t-sections of level sets of the value function are not convex too. For example, here, for the case of varying advantage of the pursuers, we see t-sections of a collection of level sets W_c at the instants $t = 12.5$ and $t = 16.0$. The t-sections are not convex and some of them even are not connected. Nevertheless here, each straight line with the direction $D_1(t)$ or $D_2(t)$ contains only one interval of points of local minimum of the restriction of the value function to this line (excluding the infinite left and right rays of constancy of the restriction). Using the directions $D_1(t)$ and $D_2(t)$, we can find the switching lines for the components of the controls of the first player.

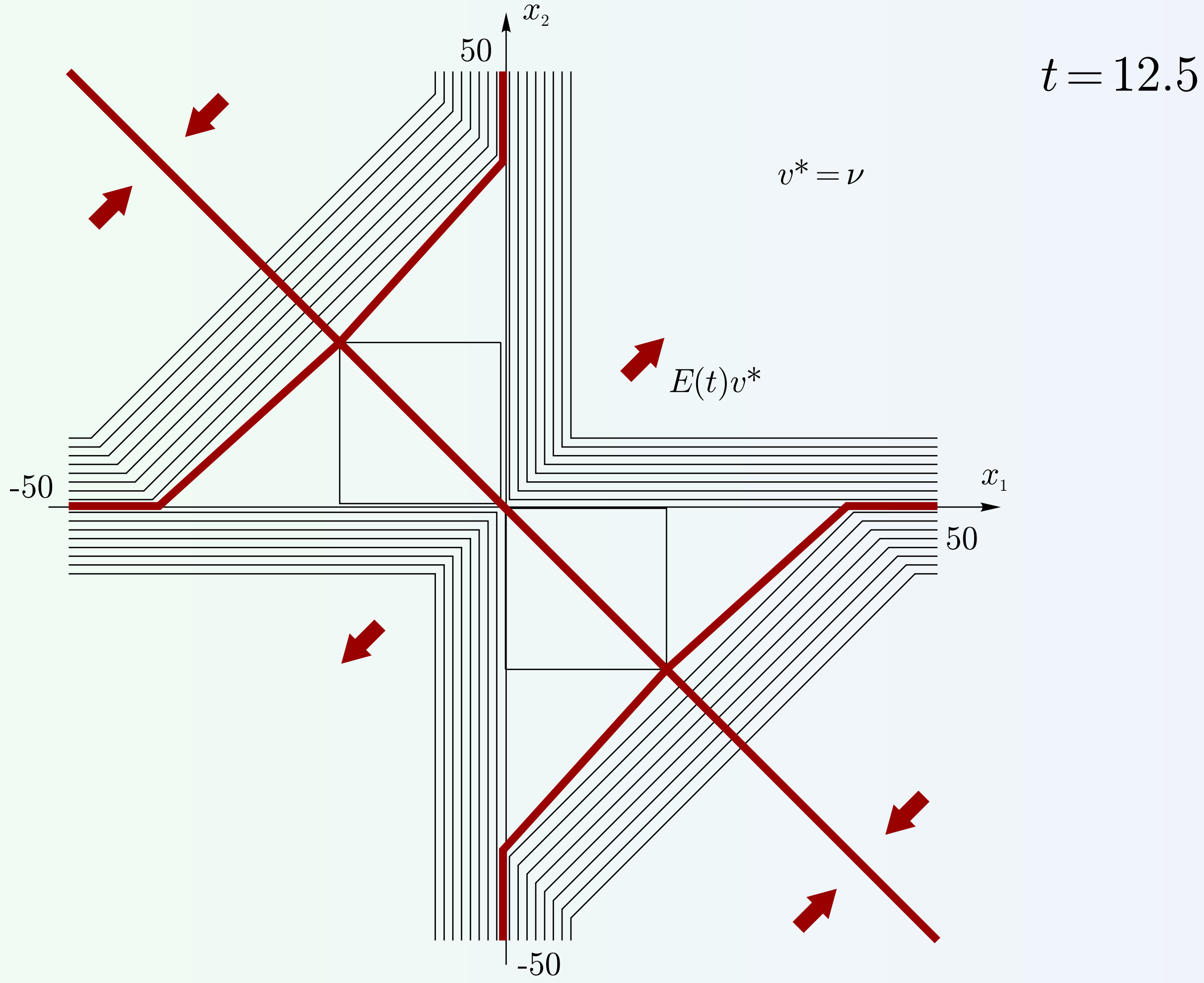
For the second player, we analyze the restrictions of the value function to straight lines parallel to the vector $E(t)$. We can see that in the pictures of this slide, we have no more than two situations of local minimum and one of local maximum. They determine the switching lines for the second player.

Switching Lines of the First Player (Varying Advantage of Pursuers, Variant 1)



In this slide, we see switching lines for optimal controls u_1^* and u_2^* of the first player.

Switching Lines of the Second Player (Varying Advantage of Pursuers, Variant 1)



Here, the switching lines of the second player are shown. For the case considered, we have six cells where the optimal control of the second player is constant and takes this or that extreme value. We have to pay attention to the directions of the arrows $E(t)v^*$. Somewhere they are directed towards the switching lines, and somewhere from the switching lines. The first case has properties similar to the one we have in the convex situation for the first player. The latter is similar to the situation of the second player in the convex case.

Computer Demonstration

The authors have elaborated an algorithm for constructing the first and second players' switching lines (depending on time) for all variants of the problem. The algorithm is based on the analysis of situations of local minima and maxima of restriction of the value function to lines parallel to vectors $D_1(t)$, $D_2(t)$, and $E(t)$.

Now, we show how do the real switching lines look and how motions of the system behave themselves under the (quasi)optimal controls based on the lines.

Reference to the video: http://home.imm.uran.ru/sector3/isdg2012/VaryingAdvantage.avi

At first, we start a program written by Sergey Ganebny for computing solution of this game: level sets of the value function, switching lines for both players, and quasioptimal control based on the switching lines. Also, this program allows to apply the computed controls and generate trajectories of the system. We can use either the optimal control for one or both players, or random control(s), or fixed constant control(s).

Let us show a modeling for the case of varying advantage of pursuers, variant 1.

We enter parameters of the game: the values η_i (in the program interface, they are denoted as "**mu1/mu2**"), ε_i, T_i (they are denoted as "**Tf1/Tf2**"). Since the pursuers are identical, their parameters coincide. Also, we enter Δ (step of the discrete scheme), upper bound of the range for the values c, and number of the values, for which the level set is computed (the grid on c is uniform).

Then, we check the option to compute a motion and enter its parameters: initial geometric deviations of the evader from the pursuers, initial relative velocities of the pursuers. Four last parameters in the program interface are added to turn one-dimensional motion along a straight line into a two-dimensional one. We add uniform longitudinal motion to the objects: **xp1/xp2** are the initial longitudinal coordinates of the pursuers, **xe** is the initial longitudinal coordinate of the evader, and **ve** is its longitudinal velocity. Longitudinal velocities of the pursuers are computed on the basis of the initial coordinates, the velocity of the evader, and termination instants T_1 and T_2.

When all parameters are defined, we press the button "Compute" to obtain the solution. During this computation, quasioptimal controls based on the switching lines are applied (this is defined by the selectors in the middle of the right panel). One can see that the computations take a few time due to efficiency of the algorithms for processing polygons in the plane.

In the informational window of the program, the following objects are drawn: time sections of the level sets (blue lines), switching lines for the first player (green lines), switching lines of the second player (red lines), the current position of the system (black circle).

Now, we start the animation. The time will change from zero to the termination instant with the step used during computations. With changing the time, all elements will change respectively. One can see as the trajectory in the ZEM-coordinates (denoted by a thin gray line) reaches one of the switching lines of the first player and then moves along it.

(continued in the next slide)

Computer Demonstration (cont.)

Let us see the motion in the geometric coordinates with longitudinal component. The graphs at the bottom of the informational window show current controls of the players. One can see as one of the pursuers changes its control in the bang-bang regime when the motion goes near the corresponding switching line. Because the geometric coordinates are natural, we cannot see the forecasted misses and can estimate the quality of the process only at the termination instant. The final misses do not vanish in this case, so there is no exact capture in this situation.

Change the initial deviations such that the initial position is still in the fourth quadrant of the ZEM-coordinate plane, but on the other side from the switching line of the second player. Recompute the trajectory of the system. Start animation for the new motion. Now, it reaches again a switching line of the first player, but in this case another one than in the previous example. Also, the motion reaches a coordinate axis, but in contrast with the previous example, further it stays in the fourth quadrant. At the termination instant, the deviations from the pursuers have equal absolute values, but opposite signs.

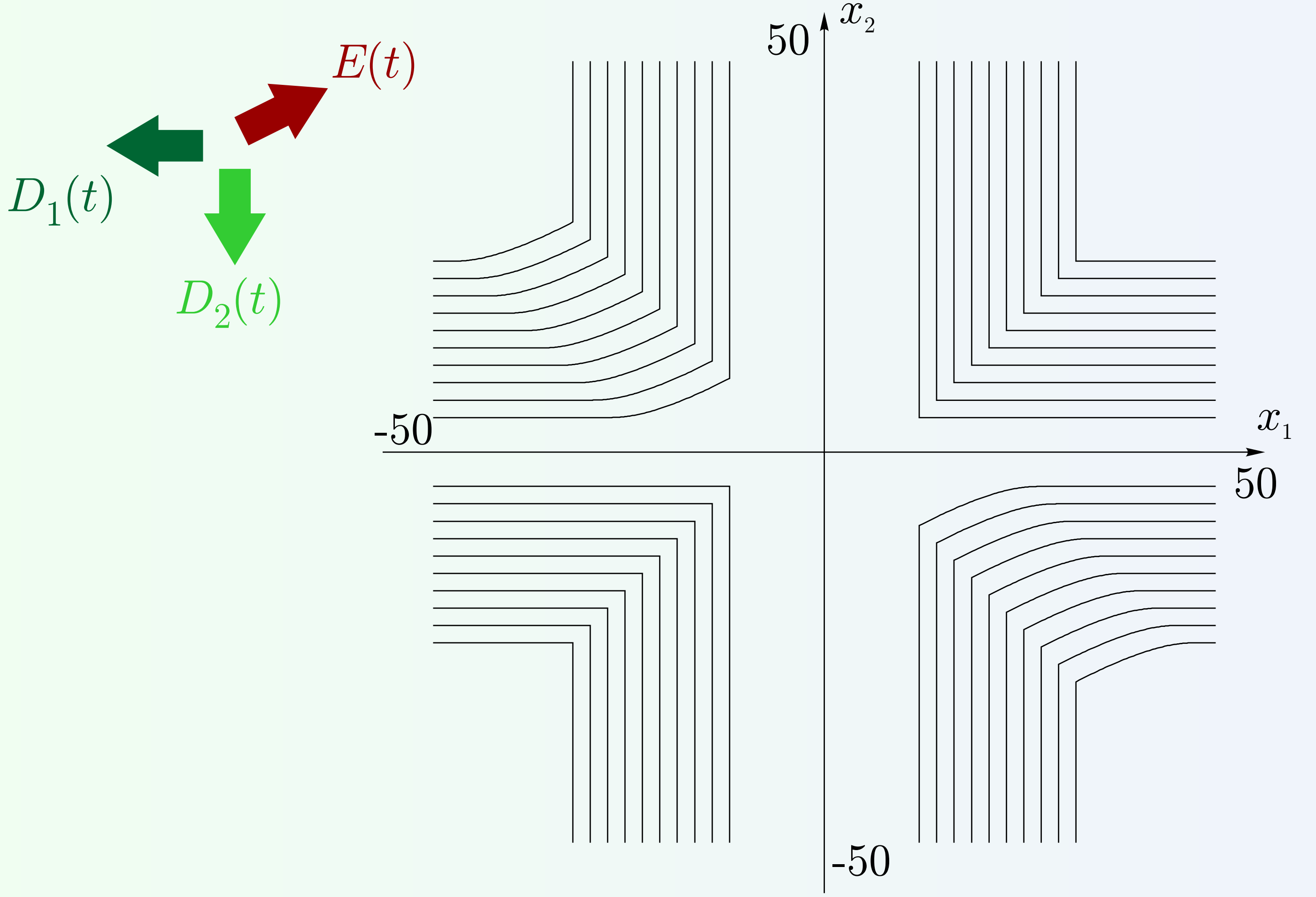
Let see the motion in the geometric coordinates. Now, the evader goes "downstairs". But it is unable to overcome the lower pursuer (the second pursuer), so, the evader aims to the middle of the final segment between the pursuers.

And finally, let us demonstrate that the controls based on the switching lines are close to the optimal ones. Change the control of the evader to the random one. "Random" means that at each instant from the time grid, the evader takes a random value from its control set $[-\nu, +\nu]$ and keeps this control value during next time interval. Recompute the trajectory. One can see that now there is no switching lines of the second player drawn in the informational window.

Under this "silly" control of the second player, the motion of the system goes to the area of small values of the game and stays there governed by the first player. Finally, the system is guided to the origin (or almost to the origin).

The motions in the original geometric space behave itself in a corresponding manner. No quick turns of the evader, because it random controls eliminate each other. The controls of the pursuers switch often as the system goes near the switching lines. And we can see that the evader is captured.

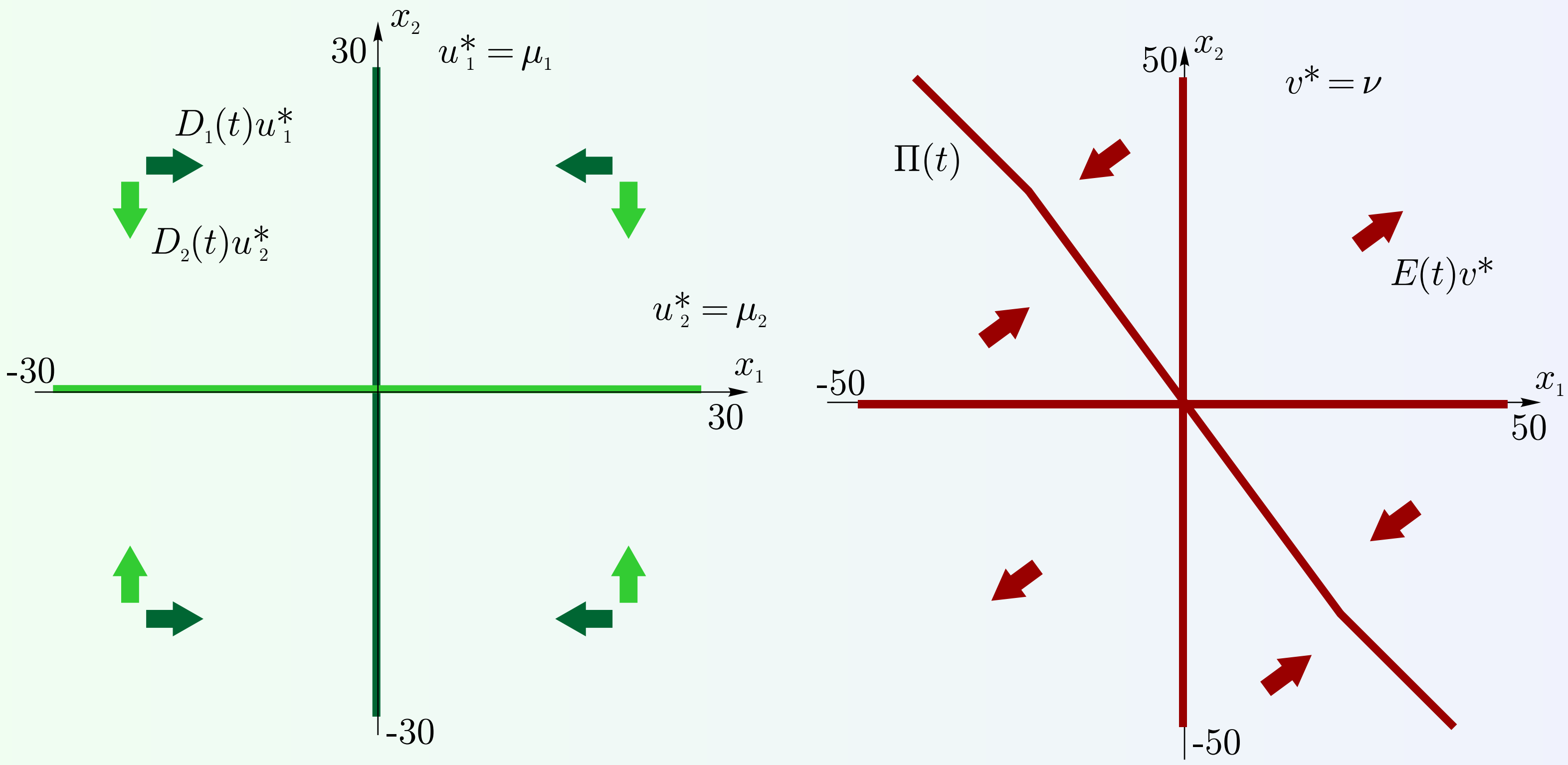
Time Sections of Level Sets (Strong Pursuers, the Case $T_1 \neq T_2$)



We have no strict mathematical proofs yet concerning the optimality of the strategies based on the switching lines. We speak about our method as quasioptimal and reasonable for practical engineering applications.

But there is a case when the proofs of the optimality are not difficult. This is the case of strong pursuers. Here, we have the following typical structure of t-sections of the level sets of the value function for some instant t.

Switching Lines
(Strong Pursuers, the Case $T_1 \neq T_2$)



The first player

The second player

The switching line for the optimal control u_1^* is the vertical axis, and the switching line for the control u_2^* is the horizontal axis. At the right, we see switching lines for the second player. Two lines are the coordinate axes, and one line is located in the second and fourth quadrant.

Case of Strong Pursuers, $T_1 \geq T_2$. Guarantees of Players

The guarantee of the first player

$$V\big(T_1, x(T_1)\big) \leq \max\big\{V\big(t_0, x(t_0)\big),\ \varepsilon + \Delta \cdot K\big\}$$

Here ε is the upper estimation of the error of system location measurement, Δ is the step of discrete control scheme of the first player, K is an estimation of the maximum of the system velocity

The guarantee of the second player

$$V\big(T_1, x(T_1)\big) \geq \max\big\{0, V\big(t_0, x(t_0)\big) - \Lambda(T_1, r, \Delta)\big\},$$

$$\Lambda(T_1, r, \Delta) = 2\sqrt{(2K\nu\Delta + r)\beta\nu} \cdot T_1 + 4K\nu\Delta + r$$

Here β is the Lipschitz constant of the function $t \to E(t)$, r is an upper estimation of the accuracy of constructing the second player's switching lines, Δ is the step of discrete control scheme of the second player. It is assumed for the second player that $\varepsilon \leq r$.

These are the upper bound for the first player guarantee and for the lower bound for the guarantee of the second player provided by the optimal control based on the switching lines in the case of strong pursuers.

References

1. Krasovskii N.N., Subbotin A.I. *Game-Theoretical Control Problems*. Springer-Verlag, New York, 1988.
2. Shinar J., Shima T. *Non-orthodox guidance law development approach for intercepting maneuvering targets*. In: Journal of Guidance, Control, and Dynamics, Vol. 25, No. 4, 2002, pp. 658-666.
3. Le Menec S. *Linear differential game with two pursuers and one evader*. In: Annals of the International Society of Dynamic Games, Vol.11: Advances in Dynamic Games. Theory, Applications, and Numerical Methods for Differential and Stochastic Games, M.Breton, K.Szajowski (Eds.), Birkhauser, Boston, 2011, pp. 209-226.
4. Ganebny S.A., Kumkov S.S., Le Ménec S., Patsko V.S. *Model problem in a line with two pursuers and one evader*. In: Dynamic Games and Applications, Vol. 2, No. 2, 2012, pp. 228-257.

In this slide, there are the main references to basic theoretical works and to works where some early investigations of this and similar problems were set forth.