REMARKS AND ERRATA*

Abstract. We provide some remarks and corrections for "Hamilton–Jacobi Formulation for Reach–Avoid Differential Games", Margellos & Lygeros, IEEE Trans. Autom. Control, 56(8), 2011, hereafter referred to as [11].

1. Victory domains. Based on [3], the following two victory domains are considered.
1. Victor's victory domain is the set ("lower" value problem): there exists a non-anticipative strategy for the control input (Victor), such that for all time measurable disturbance inputs (Ursula), the solution remains in some set K for all time instances.
2. Ursula's victory domain is the complement of the set ("upper" value problem): for all disturbance non-anticipative strategies (Ursula), for all $\epsilon > 0$, $T \geq 0$, there exists a time measurable control input (Victor), such that the solution remains in $K + \epsilon$Ball, for all time instances up to time T.

Note that in Ursula's victory domain, $K + \epsilon$Ball appears in place of K (to account for the fact that unlike Victor, Ursula's victory domain would involve an open set if K is closed). If we only consider K, Example 1 below provides an instance where there exists a non-anticipative strategy for Ursula to leave K for any choice of Victor, so Ursula's victory domain would be $\mathbb{R}^n$ while 0 belongs to Victor's victory domain (as shown in the next subsection). As a result the victory domains would no longer form a partition, and starting at 0 the player that knows the other's input (i.e., the one playing a non-anticipative strategy) could win the game.

Example 1 (taken from remark below Theorem 2.3 in [3]): Consider a vector field $f(x, u, v) = u + v$, with $u, v \in [-1, 1]$, and target set $K = \{0\}$. For the following non-anticipative strategy for Ursula there is no control input for Victor to keep the system state in $\{0\}$. Such a strategy is $\gamma[v] = v$ if $v \neq 0$, and $\gamma[v] = 1$ if $v = 0$.

2. Discriminating and leadership kernels. Discriminating and leadership kernels are defined either via inequalities satisfied by the Hamiltonian of the corresponding optimal control problem using non-smooth analysis techniques (see [3], [6] for infinite horizon problems), or via zero-sublevel sets in the sense $\{x : V(x, t) \leq 0\}$ for "lower" and "upper" value functions – we do not need their explicit characterization but the fact that these are linked to a Hamilton-Jacobi-Bellman equation [9], [12]. These value functions can be related to Victor's and the complement of Ursula's victory domain, respectively. Kernels and victory domains coincide with each other *only* if non-anticipative strategies are used. In fact, Victor's victory domain is a subset of the discriminating kernel, while Ursula's victory domain is a subset of the complement of the leadership kernel, if the corresponding player is restricted to feedback strategies [4]. Consider the following example [4], [8].

Example 2: Consider a vector field $f(x, u, v) = (1 + |x|)u + v$, with $u, v \in [-1, 1]$, and target set $K = [-1, 1]$. It is shown that under feedback strategies on (x, t), 0 is not included in the victory domain of Victor, while it is included in the discriminating kernel (see discussion below for a justification). For any feedback strategy $v(x, t)$ of Victor, the following time measurable strategy for Ursula leaves K in finite time: $u(t) = v(0, t)$ if $v(0, t) \neq 0$ and $u(t) = u_0 \neq 0$ otherwise, where u_0 is an arbitrary admissible input. The vector field $f(0, u(t), v(0, t))$ is never zero so the state trajectory escapes $\{0\}$ in finite time. After that time, setting $u(t) = x_0/\|x_0\|$ (where x_0 is the state when the trajectory escapes $\{0\}$) will guarantee leaving K.

Consider [11] for the reach-avoid variant of the problem, where the role of u and v is reversed. Note also that the value function $V(x, t) = \inf_{\gamma \in \Gamma_{[t,T]}} \sup_{v \in \mathcal{V}_{[t,T]}} \max \left\{ l(\phi(T, x, \gamma(\cdot), v(\cdot))), \max_{\tau \in [t,T]} h(\phi(\tau, x, \gamma(\cdot), v(\cdot))) \right\}$ considered in [11] is a "lower" value function. We introduce the subscript F in $\mathrm{RA}_F(t, \mathrm{R}, \mathrm{A})$ and $V_F(x, t)$, when feedback as opposed to non-anticipative strategies are used. With a proof identical to the sufficiency part of Proposition 1 in [11], however, using feedback strategies instead, we have that $\mathrm{RA}_F(t, \mathrm{R}, \mathrm{A}) \subseteq \{x : V_F(x, t) \leq 0\}$. Notice that for any (x, t), $V(x, t) \leq V_F(x, t)$, since the class of feedback strategies is restricted with respect to the one of non-anticipative strategies (feedback strategies are non-anticipative *and* playable in the sense of [2], Definition 3.1, p. 449) and the control input is minimizing in the definition of V and V_F. Hence, we have that $\{x : V_F(x, t) \leq 0\} \subseteq \{x : V(x, t) \leq 0\}$, and as a result $\mathrm{RA}_F(t, \mathrm{R}, \mathrm{A}) \subseteq \{x : V(x, t) \leq 0\}$. Note that we are interested in relating $\mathrm{RA}_F(t, \mathrm{R}, \mathrm{A})$ that depends on feedback strategies with V, that in turn depends on non-anticipative strategies, as the latter is computable as it is the (unique) solution of a quasi-variational inequality (involving an infinite number – one for each x – of finite dimensional optimization problems). This is established in Theorem 1 of [11], where non-anticipative strategies are implicitly required in the proof of Theorem 1 therein.

Example 2 offers additional insight on this issue, showing that the established inclusion under feedback strategies might be strict, namely $\mathrm{RA}_F(t, \mathrm{R}, \mathrm{A}) \subset \{x : V(x, t) \leq 0\}$. In fact, "0" might not be contained in $\mathrm{RA}_F(t, \mathrm{R}, \mathrm{A})$ (equivalent to Victor's victory domain), while it is included in $\{x : V(x, t) \leq 0\}$ (equivalent to

*20.02.2019

the discriminating kernel for reach-avoid problems). To see the latter notice that V is related to the "lower" Hamiltonian $H(p,x) = \sup_v \inf_u p^\top f(x,u,v)$ which is zero when x is zero as a result of the bang-bang nature and the symmetry of the problem. In fact, the costate vector is zero in that case, and the value function is differentiable and satisfies a Hamilton-Jacobi equation in the classical sense, as shown in [8]. At initialization (terminal condition at time instance T), "0" is included in the sublevel set of the value function, hence $V(0,T) \leq 0$. As $\frac{\partial V(0,t)}{\partial t} + H(p,0) = \frac{\partial V(0,t)}{\partial t} = 0$, we obtain that $V(0,t) = V(0,T) \leq 0$ for all $t \in [0,T]$, showing that $0 \in \{x : V(x,t) \leq 0\}$. In fact, $x = 0$ is a semipermeable barrier [3], [4] (Isaac's equation), i.e., part of the boundary of the discriminating kernel but not of the boundary of K, such that for the winning non-anticipative strategy of Victor, there exists an input for Ursula (in this example for all Ursula's inputs) so that the system trajectories stay at the boundary. To see this, notice that the non-anticipative strategy for Victor (u in the notation of [11]) is $\gamma[v] = -v$, for any admissible input v. For $x = 0$, this results in $f = 0$, hence trajectories emanating from $x = 0$ will never escape that point under the adopted strategies.

3. Value of game. A game has a value if any initial state x belongs exactly to one of the two victory domains, or in other words if the victory domains form a partition of the space [4]. Under Isaacs condition the order of the inf (control) and sup (disturbance) operators in the Hamiltonian[1] of the optimal control problem can be reversed without affecting its value. As a result, the "lower" and the "upper" value functions will coincide. The corresponding zero-sublevel sets will also be the same giving rise to a partition of the space, thus the discriminating and leadership kernels will be equal to each other. However, the victory domains are not necessarily the same with the discriminating and leadership kernels (e.g., recall that Victor's victory domain is a subset of the discriminating kernel), thus not forming a partition, unless non-anticipative strategies are used (see last paragraph of Section 2, where we may have $\mathrm{RA}_F(t,\mathrm{R},\mathrm{A}) \subset \{x : V(x,t) \leq 0\}$). Therefore, a game has a value if Isaacs condition holds *and* non-anticipative strategies are used [3].

4. Existence of a non-anticipative strategy. Consider the setting and notation of [11], where the "lower" value problem is analyzed (recall that the roles of u and v are reversed here). We point out an omission in the proof of the second part in Proposition 1 (also in Proposition 2 which follows similar arguments) of [11]. In particular, ϵ cannot be set to $\delta/2$ as δ depends on the chosen non-anticipative strategy, which in turn depends on ϵ. Therefore, rather than claiming that for all $\epsilon > 0$, there exists a non-anticipative strategy, existence of a non-anticipative strategy needs to be ensured in an alternative manner. This can be achieved by invoking the measurable viability theorem with target [1], and revising the proof according to the necessity part of the proof of Theorem 3 in [6]. To this end, we further need to assume that for all x and $v \in V$, the set $\bigcup_u f(x,u,v)$ is convex and compact (see eq. (17) in [4]); compactness follows from boundedness of f and compactness of U. Alternatively, convexity of U – Victor's domain in that notation – and affinity in u could be imposed, but the latter is a stricter condition [3]. Note that the union is over u, as we effectively require (upon the control non-anticipative strategy is fixed) that for each $v(\cdot)$, a possibly different $u(\cdot)$ exists such that the reachability specification is met. For the "upper" value problem this assumption is not required as the order of the quantifiers is reversed, and due to the presence of the ϵ-ball in the associated victory domain (see eq. (6) in [3], Assumption 3 in [6]). By Filippov's theorem [7], convexity and compactness of $\bigcup_u f(x,u,v)$ ensures that the sets reached by the trajectories $\phi(t,x,\gamma,v)$ of the dynamical system are compact; alternatively, relaxed controls should be employed (see p. 444 in [2] or [5]).

Formally, to show that $\{x : V(x,t) \leq 0\} \subseteq \mathrm{RA}(t,\mathrm{R},\mathrm{A})$, it is assumed for the sake of contradiction that $x \notin \mathrm{RA}(t,\mathrm{R},\mathrm{A})$, hence for all γ, there exists v such that $l(\phi(T,x,\gamma,v)) > 0$, where $l(\cdot)$ encodes the reach set; no need to introduce $\delta > 0$. A similar reasoning applies for the time that the system trajectories enter the avoid set, replacing $l(\cdot)$ with the function $h(\cdot)$ that encodes the avoid set. However, by the fact that $V(x,t) \leq 0$, there exists a sequence of $\epsilon_k > 0$ tending to zero as $k \to \infty$, such that for each ϵ_k, there exists γ_k, such that for all v, $l(\phi(T,x,\gamma_k,v) \leq \epsilon_k$ (equivalently, "for all v there exists u_k that gives rise to a non-anticipative strategy γ_k", hence we are interested in convexity and compactness of $\bigcup_u f(x,u,v)$). By the convexity and compactness

[1]In the definition of the Hamiltonian the order of the inf and sup operators is reversed compared to the one in the corresponding value function definition (see p. 434 in [2]). To see this, take as an example the value function $V(x,t)$ considered in [11]. The inner operator (sup) is taken with respect to time measurable functions for the disturbance, while the outer operator (inf) is with respect to non-anticipative strategies for the control (which are functions of the opponent's time measurable input); in the Hamiltonian both operators are with respect to the same class, namely the set of valuations of each player's input. Intuitively, once a non-anticipative strategy is fixed, then the disturbance makes a choice (plays first) and the control tailors its action to it (plays second) according to the chosen strategy, hence for each disturbance a possibly different control exists. Thinking of the Hamiltonian of the same problem, the inner operator is inf (control) and the outer is sup (disturbance), thus control is parametric with respect to disturbance, consistent with the order of play following the definition of $V(x,t)$. Like weak duality, $\sup\inf \leq \inf\sup$, thus justifying the fact that $\sup\inf$ in the Hamiltonian is related to the "lower" value function $V(x,t)$. As the discriminating kernel is the zero-sublevel set of $V(x,t)$, it is always a superset (or equal) to the leadership kernel, since it corresponds to a lower value function, and hence a larger sublevel set.

assumption, reachable sets are compact, hence there exists a subsequence such that $\phi(T, x, \gamma_k, v)$ converges to a system trajectory $\phi(T, x, \gamma, v)$ (corresponding to some non-anticipative strategy γ) across this subsequence. Since $l(\cdot)$ is Lipschitz continuous with Lipschitz constant $C_l > 0$, we have that

$$\begin{aligned} 0 < l(\phi(T, x, \gamma, v)) &\leq l(\phi(T, x, \gamma_k, v)) + C_l \|\phi(T, x, \gamma_k, v) - \phi(T, x, \gamma, v)\| \\ &\leq \epsilon_k + C_l \|\phi(T, x, \gamma_k, v) - \phi(T, x, \gamma, v)\|. \end{aligned} \tag{4.1}$$

Taking the limit $k \to \infty$ across the aforementioned subsequence, we have that $0 < l(\phi(T, x, \gamma, v)) \leq 0$, thus establishing a contradiction. This proof line follows the one adopted for single input reachability problems in [10].

The convexity assumption is omitted from [11] (it is only needed for the proofs of Propositions 1 and 2, while it is not required in the proofs of Theorems 1 and 2 therein), and [12], while in [9] the statement does not require such an assumption as the sets are assumed to be open and a single input setting is considered. In fact, such an assumption is not needed in reachability or invariance calculations in single input problems – it appears to be needed when having an interplay of existential and universal quantifiers as in two player games involving viability or controlled invariance (also in reach-avoid) computations.

To synthesize the players' strategies, we need to determine the optimizers of the associated Hamiltonian. These are state feedback strategies (depending on the costate vector) apart from singular cases where for the game to have a value we need to rely on knowledge of the current opponent's strategy, thus constructing a non-anticipative strategy. As an example, in the pathological situation in Example 2 where the costate vector is zero at $x = 0$, for any admissible control/disturbance $0 \in \{x : V(x, t) \leq 0\}$ (multiple optimizers exist), however, only for the optimizers corresponding to the non-anticipative strategy $0 \in \mathrm{RA}(t, \mathrm{R}, \mathrm{A})$. From a numerical point of view this is typically not a consideration as such singular points are often of measure zero, as in Example 2.

5. Correction to proof of Theorem 1 in [11]. Note that the equation above (10), pp. 1853, should be

$$\sup_{v(\cdot) \in \mathcal{V}_{[t_0, t_0+\delta_3]}} h(\phi(\tau_0, t_0, x_0, \gamma(\cdot), v(\cdot))) = \sup_{v(\cdot) \in \mathcal{V}_{[t_0, t_0+\delta_3]}} \max_{\tau \in [t_0, t_0+\delta_3]} h(\phi(\tau, t_0, x_0, \gamma(\cdot), v(\cdot))). \tag{5.1}$$

This in turn implies that $\max_{\tau \in [t_0, t_0+\delta_3]} h(\phi(\tau, t_0, x_0, \gamma(\cdot), \hat{v}(\cdot))) \neq h(\phi(\tau_0, t_0, x_0, \gamma(\cdot), \hat{v}(\cdot)))$, as τ_0 is the time where h is maximized across the worst-case $v(\cdot)$ and not only for $\hat{v}(\cdot)$. However, all statements remain valid if we maximize with respect to time τ_0 prior to fixing the strategy in the equation below (10), Case 1.1, Part 1, proof of Theorem 1 (similarly for Case 1.2). By Lemma 1, and since $\sup_{v(\cdot) \in \mathcal{V}_{[t_0, \tau_0]}} \max_{\tau \in [t_0, \tau_0]} h(\phi(\tau, t_0, x_0, \gamma(\cdot), v(\cdot))) \leq \sup_{v(\cdot) \in \mathcal{V}_{[t_0, t_0+\delta_3]}} \max_{\tau \in [t_0, t_0+\delta_3]} h(\phi(\tau, t_0, x_0, \gamma(\cdot), v(\cdot)))$ as $0 < \tau_0 \leq t_0 + \delta_3$, we have the first inequality below,

$$\begin{aligned} V(x_0, t_0) &\leq \sup_{v(\cdot) \in \mathcal{V}_{[t_0, t_0+\delta_3]}} \Big[\max \Big\{ \max_{\tau \in [t_0, t_0+\delta_3]} h(\phi(\tau, t_0, x_0, \gamma(\cdot), v(\cdot))), V(\phi(\tau_0, t_0, x_0, \gamma(\cdot), v(\cdot)), \tau_0) \Big\} \Big] \\ &= \sup_{v(\cdot) \in \mathcal{V}_{[t_0, t_0+\delta_3]}} \Big[\max \Big\{ h(\phi(\tau_0, t_0, x_0, \gamma(\cdot), v(\cdot))), V(\phi(\tau_0, t_0, x_0, \gamma(\cdot), v(\cdot)), \tau_0) \Big\} \Big] \\ &\leq \max \Big\{ h(\phi(\tau_0, t_0, x_0, \gamma(\cdot), \hat{v}(\cdot))), V(\phi(\tau_0, t_0, x_0, \gamma(\cdot), \hat{v}(\cdot)), \tau_0) \Big\} + \epsilon, \end{aligned} \tag{5.2}$$

where the eequality is by exchanging $\sup - \max$ and due to (5.1), and the last inequality is since for any $\epsilon > 0$, we can choose $\hat{v}(\cdot) \in \mathcal{V}_{[t_0, t_0+\delta_3]}$ such that the inequality holds. This issue does not appear in single-input problems.

REFERENCES

[1] J. Aubin. Viability theory. *Springer*, 2009.
[2] M. Bardi and I. Capuzzo-Dolcetta. Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. *Modern Birkhauser Classics*, 1997.
[3] P. Cardaliaguet. A differential game with two players and one target. *SIAM J. Control Optimiz.*, 34(4):1440–1460, 1996.
[4] P. Cardaliaguet. Nonsmooth, semipermeable barriers, Isaac's equation and application to a differential game with one target and two players. *Appl. Math. Optimiz.*, 34(4):125–146, 1997.
[5] I. J. Fialho and T. T. Georgiou. Worst case analysis of nonlinear systems. *IEEE Trans. Autom. Control*, 44(6):1180–1196, 1999.
[6] Y. Gao, J. Lygeros, and M. Quincampoix. On the reachability problem for uncertain hybrid systems. *IEEE Trans. Autom. Control*, 52(9):1572–1586, 2007.
[7] D. Liberzon. Calculus of variations and optimal control theory: A concise introduction. *Princeton University Press*, 2012.
[8] J. Lygeros. Worked examples of invariant set computations. *Tech. Rep.*, pages 1 – 12, 2001.
[9] J. Lygeros. On reachability and minimum cost optimal control. *Automatica*, 40(6):917 – 927, 2004.
[10] J. Lygeros, S. Sastry, and C. Tomlin. Hybrid systems: Foundations, advanced topics and applications. *Book Draft*, 2012.
[11] K. Margellos and J. Lygeros. Hamilton-Jacobi formulation for reach-avoid differential games. *IEEE Trans. Autom. Control*, 56(8):1849–1861, 2011.
[12] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin. A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games. *IEEE Trans. Autom. Control*, 50(7):947–957, 2005.