Lipschitz continuity of the solutions to team optimization problems revisited

Giorgio Gnecco and Marcello Sanguineti

Abstract—Sufficient conditions for the existence and Lipschitz continuity of optimal strategies for static team optimization problems are studied. Revised statements and proofs of some results in "Kim K.H., Roush F.W., Team Theory. Ellis Horwood Limited Publishers, Chichester, UK, 1987" are presented.

Keywords—Statistical information structure, team utility function, value of a team, Lipschitz continuity.

I. Introduction

In team optimization problems, a group of *decision makers* (*DMs*), each having at disposal some information and various possibilities of *decisions*, is interested in maximizing a common goal, expressed via a *team utility function*. Each DM takes a decision as a function, called *strategy*, of its available information. In the model that we adopt in this paper, the information is expressed via a probability density function, so we have a *statistical information structure* [3, Chapter 3]. We consider *static team optimization problems* [5], in which the information of each DM depends on a random variable, called *state of the world*, but not on the decisions of the other DMs. Otherwise, one has a *dynamic team optimization problem*; it was shown in [8] that many dynamic team optimization problems can be reformulated in terms of equivalent static ones.

Closed-form solutions for both static and dynamic team optimization problems can be derived only under quite strong assumptions on the team utility function and the way in which each DM's information is influenced by the state of the world (and, in the case of dynamic teams, by the decisions of the other DMs) [3]. If these conditions are not met, one has to search for approximate solutions. In such a case, knowing structural properties of optimal strategies (e.g., Lipschitz continuity) is useful to find good suboptimal strategies.

The aim of this paper is to present revised statements and proofs of some results appeared in [3, Section 5.2] on existence and Lipschitz continuity of optimal strategies for a family of static team optimization problems. Although the book [3] offers an interesting and inspirational exposition of the mathematical theory of team optimization problems, it contains various misprints, omissions, and technical inconsistencies, probably due to a too fast and inaccurate writing. These drawbacks were pointed out from the very beginning in a couple of reviews [4], [6]. We quote from [6]: "The strength of the book lies in the power and originality of the ideas used to achieve its stated goal of extending the theory of teams in a number of new directions" but "Unfortunately, the book is beset with a variety of technical problems that will prevent all but the most tolerant, persistent, and experienced readers from reaping the benefits of the later chapters". We hope that our work will make the revisited results more easily accessible and usable.

The paper is organized as follows. Section II introduces definitions and assumptions and formulates the family of static team optimization problems under consideration. Section III presents revised statements and proofs of some results appeared in [3].

Giorgio Gnecco is with the Department of Communications, Computer and System Sciences (DIST) and the Department of Computer and Information Science (DISI), University of Genoa, Italy (phone: +39 010 353 2071-2983; fax: +39 010 3532154; e-mail: giorgio.gnecco@dist.unige.it).

Marcello Sanguineti is with the Department of Communications, Computer and System Sciences (DIST), University of Genoa, Italy (e-mail: marcello@dist.unige.it).

Manuscript received June 28, 2009.

II. Problem formulation

The context in which we shall formalize the optimization problem and state the results is the following.

- *Static team* of n *decision makers* (*DMs*), $i = 1, \ldots, n$.
- $x \in X \subseteq \mathbb{R}^{d_0}$: vector-valued random variable, called *state of the world*, describing a stochastic environment. The vector x models the uncertainties in the external world, which are not controlled by the DMs.
- $y_i \in Y_i \subseteq \mathbb{R}^{d_i}$: vector-valued random variable, which represents the *information* that the DM i has about x.
- $s_i : Y_i \to A_i \subseteq \mathbb{R}$: Borel-measurable *strategy* of the i-th DM.
- $a_i = s_i(y_i)$: *decision* that the DM i chooses on the basis of the information y_i.
- $u : X \times \Pi_{i=1}^n Y_i \times \Pi_{i=1}^n A_i \subseteq \mathbb{R}^N \to \mathbb{R}$, where $N = \sum_{i=0}^n d_i + n$: real-valued *team utility* function.
- The information that the n DMs have on the state of the world x is modelled by an n-tuple of random variables $y_1, \ldots, y_n$, i.e., by a *statistical information structure* [3, Chapter 3] represented by a joint probability density $q(x, y_1, \ldots, y_n)$ on the set $X \times \Pi_{i=1}^n Y_i$.

We formulate the following static team optimization problem.

Problem STO (Static Team Optimization with Statistical Information). *Given the statistical information structure* $q(x, y_1, \ldots, y_n)$ *and the team utility function* $u(x, y_1, \ldots, y_n, a_1, \ldots, a_n)$*, find*

$$\sup_{s_1,\dots s_n} v(s_1,\dots,s_n)\,,$$

where

$$v(s_1,\dots,s_n)=\mathbb{E}_{x,y_1,\dots,y_n}\left\{u(x,\{y_i\}_{i=1}^n,\{s_i(y_i)\}_{i=1}^n)\right\}.$$

The quantity $\sup_{s_1,\dots s_n} v(s_1,\dots,s_n)$ is called the *value of the team*.

Throughout the paper, we make the following two assumptions.

A1 *The set X of the states of the world is compact, $Y_1,\dots,Y_n$ are compact and convex, and $A_1,\dots,A_n$ are bounded closed intervals. The team utility u is of class $\mathcal{C}^2$ on an open set containing $X\times\Pi_{i=1}^n Y_i\times\Pi_{i=1}^n A_i$ and q is a (strictly) positive probability density on $X\times\Pi_{i=1}^n Y_i$, which can be extended to a function of class $\mathcal{C}^2$ on an open set containing $X\times\Pi_{i=1}^n Y_i$.*

For $\tau>0$, a concave function f defined on a convex set $\Omega\subseteq\mathbb{R}^d$ has *concavity (at least)* τ if for all $u,v\in\Omega$ and every supergradient[1] p_u of f at u one has $f(v)-f(u)\le p_u\cdot(v-u)-\tau\|v-u\|^2$. If f is of class $\mathcal{C}^2(\Omega)$, then a necessary condition for its concavity τ is $\sup_{u\in\Omega}\lambda_{\max}(\nabla^2 f(u))\le-\tau$, where $\lambda_{\max}(\nabla^2 f(u))$ is the maximum eigenvalue of the Hessian $\nabla^2 f(u)$. Indeed,

$$f(v)-f(u)\le p_u\cdot(v-u)-\tau\|v-u\|^2$$

implies

$$\begin{aligned}
& f(v)+\frac{1}{2}\tau\|v\|^2-f(u)-\frac{1}{2}\tau\|u\|^2\\
\le\ & p_u\cdot(v-u)-\tau\|v-u\|^2+\frac{1}{2}\tau\|v\|^2-\frac{1}{2}\tau\|u\|^2\\
=\ & (p_u+\tau u)\cdot(v-u)-\frac{\tau}{2}\|v-u\|^2\\
\le\ & (p_u+\tau u)\cdot(v-u)\,,
\end{aligned}$$

i.e., $f(\cdot)+\frac{1}{2}\tau\|\cdot\|^2$ is concave, then one applies the characterization of concavity for a function of class $\mathcal{C}^2$.

A2 *There exists $\tau>0$ such that the team utility function $u: X\times\Pi_{i=1}^n Y_i\times\Pi_{i=1}^n A_i\to\mathbb{R}$ is separately concave with concavity τ in each of the decision variables*[2].

Assumption A2 is motivated by tractability reasons and encountered in practice. For example, in economic problems it is motivated by the "law of diminishing returns", i.e., the fact that the marginal productivity of an input usually diminishes as the amount of output increases [5, p. 99 and p. 110].

[1]For $\Omega\subseteq\mathbb{R}^d$ convex and $f:\Omega\to\mathbb{R}$ concave, $p_u\in\mathbb{R}^d$ is a *supergradient* of f at $u\in\Omega$ if for every $v\in\Omega$ it satisfies $f(v)-f(u)\le p_u\cdot(v-u)$.

[2]I.e., if all the arguments of u are fixed except the decision variable a_i, then the resulting function of a_i has concavity τ.

III. Lipschitz continuity of the optimal strategies

In this section, we shall give conditions that guarantee existence and Lipschitz continuity of optimal strategies for Problem STO.

The next lemma is obtained making various changes and corrections to [3, Lemma 10, p. 162].

Lemma 1: Let $q(\gamma)$ be a probability density for the real vector-valued random variable γ with values in $\Gamma\subseteq\mathbb{R}^{m_\gamma}$, $Z=\mathbb{R}^m$ or Z a compact subset of $\mathbb{R}^m$, and $\{f_\gamma\}$ a set of functions $f_\gamma: Z\to\mathbb{R}$, parameterized by γ, equiLipschitz with constant L and concavity τ. If for every $z\in Z$ the function $f_\gamma(z):\Gamma\to\mathbb{R}$ is Borel-measurable, then the function defined for every $z\in Z$ as $f(z)=\int_\Gamma q(\gamma)f_\gamma(z)d\gamma$ is Lipschitz with constant L and concavity τ.

Proof. Lipschitz continuity with constant L follows by

$$\begin{aligned}
|f(z)-f(w)| &= \left|\int_\Gamma q(\gamma)[f_\gamma(z)-f_\gamma(w)]d\gamma\right|\\
&\le \int_\Gamma q(\gamma)L\|z-w\|d\gamma=L\|z-w\|\,.
\end{aligned}$$

Let us prove the statement about concavity. By assumption, for every $\gamma\in\Gamma$ we have

$$f_\gamma(z)-f_\gamma(w)\le a_\gamma(w)\cdot(z-w)-\tau\|z-w\|^2\,, \tag{1}$$

where $a_\gamma(w)$ is a supergradient of f_γ at w.

By [2, Proposition 2.2.7, p. 36 and Theorem 2.7.2, p. 76] (which can be applied since, if $Z=\mathbb{R}^m$ or is a compact subset of $\mathbb{R}^m$, then Z is separable), every supergradient $a(w)$ of $f(z)=\int_\Gamma q(\gamma)f_\gamma(z)\,d\gamma$ at w can be written in the form

$$a(w)=\int_\Gamma q(\gamma)a_\gamma^{\mathrm{m}}(w)d\gamma\,, \tag{2}$$

where $a_\gamma^{\mathrm{m}}(w)$ is a measurable selection (with respect to γ) of the set $\partial f_\gamma(w)$ of all supergradients of f_γ at w. With such a choice of the supergradient, by taking expectations in (1) and using (2) we get

$$\begin{aligned}
f(z)-f(w) &= \int_\Gamma q(\gamma)[f_\gamma(z)-f_\gamma(w)]d\gamma\\
&\le \int_\Gamma q(\gamma)a_\gamma^{\mathrm{m}}(w)d\gamma\cdot(z-w)-\tau\|z-w\|^2\\
&= a(w)\cdot(z-w)-\tau\|z-w\|^2\,,
\end{aligned}$$

i.e., f has concavity τ. ■

In the proof of Theorem 1, we shall exploit the following known result, which for completeness we report here together with its proof.

Lemma 2: Let Z be a subset of a normed linear space, $\{f_k\}$ a sequence of real-valued functions on Z, equiLipschitz with constant L. If for every $z \in Z$ their point-wise limit $f(z) = \lim_{k\to\infty} f_k(z)$ exists, then f is Lipschitz with constant L.

Proof. By hypothesis, for every $x, y \in Z$ we have $|f_k(y) - f_k(x)| \le L\|y-x\|$. Then, $\lim_{k\to\infty} |f_k(y) - f_k(x)| = |f(y) - f(x)| \le L\|y-x\|$. ■

The following theorem, obtained making various changes and corrections to [3, Theorem 11, p. 162], provides conditions guaranteeing that Problem STO has a solution made of an n-tuple of Lipschitz strategies.

Theorem 1: Under Assumptions A1 and A2, Problem STO admits Lipschitz optimal strategies.

Proof. We detail the proof for the case of $n = 2$ DMs, then we mention the changes required for the extension to $n > 2$.

Proof for $n = 2$.

Consider a sequence $\{s_1^j, s_2^j\}$ of pairs of strategies, indexed by $j \in \mathbb{N}_+$, such that

$$\lim_{j\to\infty} v(s_1^j, s_2^j) = \sup_{s_1, s_2} v(s_1, s_2)$$

(such a sequence exists by the definition of supremum). From this sequence, we generate another sequence $\{\hat{s}_1^j, \hat{s}_2^j\}$ defined for every $y_1 \in Y_1$ and every $y_2 \in Y_2$ as

$$\hat{s}_1^j(y_1) = \underset{a_1\in A_1}{\operatorname{argmax}}\, \mathbb{E}_{x,y_2\,|y_1} \{u(x, y_1, y_2, a_1, s_2^j(y_2))\}\,, \quad (3)$$

$$\hat{s}_2^j(y_2) = \underset{a_2\in A_2}{\operatorname{argmax}}\, \mathbb{E}_{x,y_1\,|y_2} \{u(x, y_1, y_2, \hat{s}_1^j(y_1), a_2)\}\,. \quad (4)$$

The proof is structured in the following steps.

Step 1. For every $j \in \mathbb{N}_+$, $\hat{s}_1^j$ and $\hat{s}_2^j$ are well-defined (i.e., the maxima in (3) and (4) exist and are unique) and Borel-measurable, so it makes sense to evaluate $v(\hat{s}_1^j, \hat{s}_2^j)$. By construction, $v(\hat{s}_1^j, \hat{s}_2^j) \ge v(s_1^j, s_2^j)$, then $\lim_{j\to\infty} v(\hat{s}_1^j, \hat{s}_2^j) = \sup_{s_1,s_2} v(s_1, s_2)$.

Step 2. For every $j \in \mathbb{N}_+$, the functions $\hat{s}_1^j$ and $\hat{s}_2^j$ are Lipschitz, with a constant independent of j.

Step 3. For every $j \in \mathbb{N}_+$, the functions $\hat{s}_1^j$ and $\hat{s}_2^j$ are equibounded and uniformly equicontinuous, so we can apply Ascoli-Arzelà's theorem [1, Theorem 1.30, p. 10] to obtain convergence of a subsequence to a pair of continuous strategies $\{s_1^o, s_2^o\}$.

Step 4. We exploit Lemma 2 and continuity of the functional $v(s_1, s_2)$ to show that the pair of strategies $\{s_1^o, s_2^o\}$ is Lipschitz and optimal.

Step 1. We make the proof for $\hat{s}_1^j$; the same arguments hold for $\hat{s}_2^j$. Let us show that for every $j \in \mathbb{N}_+$ the functions $\hat{s}_1^j$ are well-defined and continuous, hence Borel-measurable. Let

$$M_1^j(y_1, a_1) = \mathbb{E}_{x,y_2|y_1} \{u(x, y_1, y_2, a_1, s_2^j(y_2))\}\,.$$

By the definition,

$$\hat{s}_1^j(y_1) = \underset{a_1\in A_1}{\operatorname{argmax}}\, M_1^j(y_1, a_1)\,. \quad (5)$$

As the probability density $q(x, y_1, y_2)$ is continuous and strictly positive on an open set containing $X \times Y_1 \times Y_2$, the conditional density $q(x, y_2|y_1)$ is continuous on $X \times Y_1 \times Y_2$. Since $q(x, y_2|y_1)$ and u are continuous on compact sets, they are uniformly continuous. So M_1^j, as an integral dependent on parameters, is continuous on the compact set $Y_1 \times A_1$. As u is of class $\mathcal{C}^1$ on a compact set, it is Lipschitz continuous thereon, too. Let L be an upper bound on its Lipschitz constant. For every y_1, by Lemma 1 M_1^j is Lipschitz in the second variable a_1 with constant L, and has concavity τ in a_1.

By the continuity and concavity properties of M_1^j with respect to a_1, the maximum in (5) exists and is unique, so $\hat{s}_1^j$ is well-defined. Let y'_1, $y''_1 \in Y_1$. By the definition of $\hat{s}_1^j$, exploiting the concavity τ of M_1^j with respect to a_1 and taking the supergradient 0 of M_1^j with respect to the second variable at $(y'_1, \hat{s}_1^j(y'_1))$ and $(y''_1, \hat{s}_1^j(y''_1))$, respectively, we get

$$\begin{aligned} & M_1^j(y'_1, \hat{s}_1^j(y''_1)) - M_1^j(y'_1, \hat{s}_1^j(y'_1)) \\ \le\ & -\tau|\hat{s}_1^j(y''_1) - \hat{s}_1^j(y'_1)|^2 \end{aligned} \quad (6)$$

and

$$\begin{aligned} & M_1^j(y''_1, \hat{s}_1^j(y'_1)) - M_1^j(y''_1, \hat{s}_1^j(y''_1)) \\ \le\ & -\tau|\hat{s}_1^j(y'_1) - \hat{s}_1^j(y''_1)|^2\,. \end{aligned} \quad (7)$$

By (6) and (7) we get

$$\begin{aligned} & |M_1^j(y'_1, \hat{s}_1^j(y''_1)) - M_1^j(y'_1, \hat{s}_1^j(y'_1))| \\ +\ & |M_1^j(y''_1, \hat{s}_1^j(y'_1)) - M_1^j(y''_1, \hat{s}_1^j(y''_1))| \\ \ge\ & 2\tau|\hat{s}_1^j(y''_1) - \hat{s}_1^j(y'_1)|^2\,. \end{aligned} \quad (8)$$

By (8) we obtain

$$|\hat{s}_1^j(y''_1) - \hat{s}_1^j(y'_1)| \le \sqrt{\frac{L}{\tau}}\sqrt{\|y''_1 - y'_1\|}\,, \quad (9)$$

which proves the Hölder continuity of $\hat{s}_1^j$, hence its continuity. Continuity of $\hat{s}_2^j$ can be proved in the same way.

Step 2. Let us prove that $\hat{s}_1^j$ and $\hat{s}_2^j$ are Lipschitz with a Lipschitz constant independent of j. We make the proof for $\hat{s}_1^j$; the same arguments hold for $\hat{s}_2^j$. To this end, as Y_1 is convex it is sufficient to prove that the restriction of $\hat{s}_1^j$ to each line joining every two points y'_1 and y''_1 is Lipschitz, with a constant that depends neither on j, nor on the line. Consider the function $\hat{s}_1^j(y_1(t))$, where $y_1(t) = y'_1 + t(y''_1 - y'_1)$ and $0 \le t \le 1$. There are two possible cases:

1) either $\hat{s}_1^j(y_1(t))$, for all $0 \le t \le 1$, is interior to $A_1 = [a_1^l, a_1^u]$,
2) or there exists $\tilde{t} \in [0,1]$ such that $\hat{s}_1^j(y_1(\tilde{t}))$ is one of the two extremes a_1^l, a_1^u of A_1.

Case 1. When $\hat{s}_1^j(y_1) = \mathrm{argmax}_{a_1 \in A_1} M_1^j(y_1, a_1)$ is an interior point of $A_1 = [a_1^l, a_1^u]$, we have

$$\frac{\partial M_1^j}{\partial a_1}\Big|_{a_1 = \hat{s}_1^j(y_1(t))} = 0\,. \tag{10}$$

As M_1^j is of class $\mathcal{C}^2$ and has concavity τ in a_1, $\frac{\partial^2 M_1^j}{\partial {a_1}^2} \le -\tau < 0$, then we can apply the implicit function theorem to the function

$$\sigma_1(t) = \hat{s}_1^j(y_1(t))\,.$$

Taking the total derivative with respect to t of both sides of (10), we get

$$\sum_{k=1}^{d_1}\left(\frac{\partial^2 M_1^j}{\partial {\sigma_1}^2}\frac{\partial \sigma_1}{\partial y_{1,k}}\frac{\partial y_{1,k}}{\partial t} + \frac{\partial^2 M_1^j}{\partial \sigma_1 \partial y_{1,k}}\frac{\partial y_{1,k}}{\partial t}\right) =$$

$$= \frac{\partial^2 M_1^j}{\partial {\sigma_1}^2}\sum_{k=1}^{d_1}\frac{\partial \sigma_1}{\partial y_{1,k}}\frac{\partial y_{1,k}}{\partial t} + \sum_{k=1}^{d_1}\frac{\partial^2 M_1^j}{\partial \sigma_1 \partial y_{1,k}}\frac{\partial y_{1,k}}{\partial t} = 0\,. \tag{11}$$

Then $\sigma_1(t)$ is locally differentiable and by (11) we have

$$\begin{aligned}\frac{d\sigma_1(t)}{dt} &= \sum_{k=1}^{d_1}\frac{\partial \sigma_1}{\partial y_{1,k}}\frac{\partial y_{1,k}}{\partial t} = \ldots = \\ &= -\left(\frac{\partial^2 M_1^j}{\partial {\sigma_1}^2}\right)^{-1}\sum_{k=1}^{d_1}\frac{\partial^2 M_1^j}{\partial \sigma_1 \partial y_{1,k}}(y''_{1,k} - y'_{1,k})\,.\end{aligned}$$

As $\left|\frac{\partial^2 M_1^j}{\partial {\sigma_1}^2}\right|^{-1} \le \frac{1}{\tau}$ and $|y''_{1,k} - y'_{1,k}| \le \mathrm{diameter}\,(Y_1)$, it remains to find for every k an upper bound on $\left|\frac{\partial^2 M_1^j}{\partial \sigma_1 \partial y_{1,k}}\right|$ in (12), independent of y_1 and j. By the definition,

$$M_1^j(y_1, a_1) = \frac{\int_{X \times Y_2} q(x, y_1, y_2) u(x, y_1, y_2, a_1, s_2^j(y_2)) dx dy_2}{\int_{X \times Y_2} q(x, y_1, y_2) dx dy_2}\,.$$

Some elementary calculations allow to express $\frac{\partial^2 M_1^j}{\partial a_1 \partial y_{1,k}}$ as a ratio whose numerator is a polynomial in

$$\int_{X \times Y_2} \frac{\partial^i [q(x, y_1, y_2) u(x, y_1, y_2, a_1, s_2^j(y_2))]}{\partial {a_1}^a \partial {y_{1,k}}^b} dx dy_2$$

and

$$\int_{X \times Y_2} \frac{\partial^i q(x, y_1, y_2)}{\partial {a_1}^a \partial {y_{1,k}}^b} dx dy_2$$

for $i = 0, 1, 2$, $a + b = i$, whereas its denominator is

$$\left(\int_{X \times Y_2} q(x, y_1, y_2) dx dy_2\right)^3 \ge \delta > 0\,,$$

where δ is a positive constant (indepedent of y_1), whose existence and independence from y_1 are guaranteed by the hypothesis $q(x, y_1, y_2) > 0$ and the continuity of $q(x, y_1, y_2)$ on the compact set $X \times Y_1 \times Y_2$. Note that the change of order between expectation and up-to-second-order partial derivatives is justified by the fact that $q(x, y_1, y_2)$ and $u(x, y_1, y_2, a_1, a_2)$ are of class $\mathcal{C}^2$ on compact sets.

Then an upper bound on $\left|\frac{\partial^2 M_1^j}{\partial a_1 \partial y_{1,k}}\right|$ can be expressed in terms of the finite quantities

$$\sup_{y_1 \in Y_1} \int_{X \times Y_2} \sup_{a_2 \in A_2} \left|\frac{\partial^i [q(x, y_1, y_2) u(x, y_1, y_2, a_1, a_2)]}{\partial {a_1}^a \partial {y_{1,k}}^b}\right| dx dy_2$$

and

$$\sup_{y_1 \in Y_1} \int_{X \times Y_2} \sup_{a_2 \in A_2} \left|\frac{\partial^i q(x, y_1, y_2)}{\partial {a_1}^a \partial {y_{1,k}}^b}\right| dx dy_2\,,$$

where measurability of the integrands follows by [7, Property (c), p. 38]. This bound does not depend on y_1. Moreover, it does not depend on the particular choice of $s_2^j(y_2)$, so it is also independent of j.

Summing up, we obtain an upper bound independent of y_1 and j on $\left|\frac{d\,\sigma_1(t)}{d\,t}\right|$.

Case 2. We now consider the case in which there exists $\tilde{t} \in [0,1]$ such that $\hat{s}_1^j(y_1(\tilde{t}))$ is one of the two extremes a_1^l, a_1^u of A_1. Suppose, e.g., that $\hat{s}_1^j(y_1(\tilde{t})) = a_1^l$. The situation $\hat{s}_1^j(y_1(\tilde{t})) = a_1^u$ can be studied in the same way. We can limit the analysis to the case in which $y_1(\tilde{t})$ does not belong to the boundary of Y_1, which has d_1-dimensional measure equal to 0, due to the convexity and boundedness of Y_1. There are two possible subcases.

- *Subcase 1*: there exists a neighbourhood of $\tilde{t}$ such that $y_1(t) = a_1^l$ does never hold (except for $t = \tilde{t}$). Then one has

$$\lim_{t \to \tilde{t}^-} \frac{d\hat{s}_1^j(y_1(t))}{dt} = \lim_{t \to \tilde{t}^+} \frac{d\hat{s}_1^j(y_1(t))}{dt} = 0\,,$$

  hence $\frac{d\hat{s}_1^j(y_1(t))}{dt}|_{t=\tilde{t}} = 0$. Indeed, $\hat{s}_1^j$ is continuously differentiable with derivative (12) when the maximum is interior to A_1 and the limit is 0 (as the maximum is not allowed to be outside A_1);
- *Subcase 2*: there exists a non-constant sequence $\{\bar{t}_l\}$ such that $\lim_{l \to \infty} \bar{t}_l = \tilde{t}$ and $y_1(\bar{t}_l) = a_1^l, \forall l \in \mathbb{N}_+$. In general, this does not allow one to deduce the existence of $\frac{d\hat{s}_1^j(y_1(t))}{dt}|_{t=\tilde{t}}$. However, if one considers the incremental ratio $\frac{\hat{s}_1^j(y_1(t)) - \hat{s}_1^j(y_1(\tilde{t}))}{t - \tilde{t}}$ and any sequence $\{t_l\}$ (which can be different from $\{\bar{t}_l\}$) such that $\lim_{l \to \infty} t_l = \tilde{t}$, then the lim sup on $\{t_l\}$ of the absolute value of that incremental ratio is bounded from above by a constant independent of $\tilde{t}$ and j (this can be easily proved by using the results of Case 1).

Case 1 and Case 2 together imply that $\hat{s}_1^j$ is Lipschitz with a constant independent of j.

Step 3. The functions belonging to $\{\hat{s}_1^j\}$ and $\{\hat{s}_2^j\}$ are equibounded, as A_1 and A_2 are bounded intervals, and uniformly equicontinuous, thanks to the uniform bound on their Lipschitz constants. Then, by Ascoli-Arzelà's theorem, there exists a subsequence of $\{\hat{s}_1^j, \hat{s}_2^j\}$ that converges uniformly to a pair of countinuous strategies $\{s_1^o, s_2^o\}$ on the compact set $Y_1 \times Y_2$.

Step 4. By Lemma 2 the limit strategies $\{s_1^o, s_2^o\}$ are Lipschitz, with the same bound on their Lipschitz constants. Since the functional

$$v(s_1, s_2) = \mathbb{E}_{x,y_1,y_2}\{u(x, y_1, y_2, s_1(y_1), s_2(y_2))\}$$

is continuous for $s_1 \in \mathcal{C}(Y_1)$ and $s_2 \in \mathcal{C}(Y_2)$ with the respective maximum norms, we finally obtain

$$v(s_1^o, s_2^o) = \lim_{j\to\infty} v(\hat{s}_1^j, \hat{s}_2^j) = \sup_{s_1,s_2} v(s_1, s_2)\,.$$

■

Extension to $n \geq 2$.

The only significant change in the proof consists in defining as follows the n-tuple $\hat{s}_1^j, \ldots, \hat{s}_n^j$ of strategies:

$$\hat{s}_1^j(y_1) = \operatorname*{argmax}_{a_1\in A_1} \mathbb{E}_{x,\{y_i\}_{i\neq 1}\,|y_1}\{u(x, \{y_i\}_{i=1}^n, a_1, \{s_i^j(y_i)\}_{i=2}^n)\}\,,$$

$$\hat{s}_2^j(y_2) = \operatorname*{argmax}_{a_2\in A_2} \mathbb{E}_{x,\{y_i\}_{i\neq 2}\,|y_2}\{u(x, \{y_i\}_{i=1}^n, \hat{s}_1^j(y_1), a_2, \{s_i^j(y_i)\}_{i=3}^n)\})\}\,,$$

$$\ldots$$

$$\hat{s}_n^j(y_n) = \operatorname*{argmax}_{a_n\in A_n} \mathbb{E}_{x,\{y_i\}_{i\neq n}\,|y_n}\{u(x, \{y_i\}_{i=1}^n, \{\hat{s}_i^j(y_i)\}_{i=1}^{n-1}, a_n)\}\,.$$

■

References

[1] R. A. Adams. *Sobolev Spaces*. Academic Press, New York, 1975.

[2] F. H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley-Interscience, New York, 1983.

[3] K. H. Kim and F. W. Roush. *Team Theory*. Ellis Horwood Limited, Chichester, 1987.

[4] J. H. Klein. Review of 'Team Theory, by K. H. Kim and F. W. Roush, Ellis Horwood Limited (Div. of John Wiley); Halsted Press, New York, 1987' . *J. of the Operational Research Society*, 39:695–696, 1988.

[5] J. Marschak and R. Radner. *Economic Theory of Teams*. Yale University Press, New Haven, 1972.

[6] E. W. Packel. Review of 'Team Theory, by K. H. Kim and F. W. Roush, Ellis Horwood Limited (Div. of John Wiley); Halsted Press, New York, 1987' . *SIAM Review*, 30:676–677, 1988.

[7] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, Singapore, 1987.

[8] H. S. Witsenhausen. Equivalent stochastic control problems. *Mathematics of Control, Signals, and Systems*, 1:3–11, 1988.