A note on the equilibrium payoff set in stochastic games*

Johannes Hörner[†], Satoru Takahashi[‡] and Nicolas Vieille[§]

August 3, 2011

Abstract

This note provides a dual characterization of the set of perfect public equilibrium payoffs in stochastic games. As a corollary, the folk theorems of Fudenberg, Levine and Maskin (1994), Kandori and Matsushima (1998) and Hörner, Sugaya, Takahashi and Vieille (2011) obtain.

Keywords: stochastic games, repeated games, folk theorem.

JEL codes: C72, C73

1 Result

This note describes another characterization of the limit payoff set in stochastic games, or more, precisely, another characterization of the nonlinear programs whose solution is key to the description of this payoff set. See Hörner, Sugaya, Takahashi and Vieille (2011, hereafter HSTV) for notation and assumptions.

The main element of the characterization of HSTV is the solution to the following non-linear program, where $\lambda \in \mathbf{R}^I$ is fixed. Given a state $s \in S$ and a map $x : S \times Y \to \mathbf{R}^{S \times I}$, we denote by $\Gamma(s, x)$ the one-shot game with action sets A^i and payoff function

$$r(s, a_s) + \sum_{t \in S} \sum_{y \in Y} p(t, y | s, a_s) x_t(s, y),$$

*We thank David Rahman for helpful discussions. J. Hörner gratefully acknowledges financial support from NSF Grant SES 092098. N. Vieille gratefully acknowledges financial support from Fondation HEC.

[†]Yale University, 30 Hillhouse Ave., New Haven, CT 06520, USA, johannes.horner@yale.edu.

[‡]Princeton University, Fisher Hall, Princeton, NJ 08544, USA, satorut@princeton.edu.

[§]HEC Paris, 78351 Jouy-en-Josas, France, vieille@hec.fr.

where $x_t(s,y) \in \mathbf{R}^I$ is the t-th component of $x(s,y)$.

Given $\lambda \in \mathbf{R}^I$, we denote by $\mathcal{P}(\lambda)$ the maximization program

$$\sup_{v,x,\alpha} \lambda \cdot v,$$

where the supremum is taken over all $v \in \mathbf{R}^I$, $x : S \times Y \to \mathbf{R}^{S\times I}$, and $\alpha = (\alpha_s) \in (\times_{i\in I}\Delta(A^i))^S$ such that

(i) For each s, α_s is a Nash equilibrium with payoff v of the game $\Gamma(s,x)$;

(ii) For each $T \subseteq S$, for each permutation $\varphi : T \to T$ and each map $\psi : T \to Y$, one has $\lambda \cdot \sum_{s\in T} x_{\varphi(s)}(s,\psi(s)) \leq 0$.

The program $\mathcal{P}(\lambda)$ is a generalization to stochastic games of the program introduced by Fudenberg and Levine (1994) for repeated games, based in turn on the recursive representation of the payoff set given by Abreu, Pearce and Stacchetti (1990).

HSTV prove that, under some assumptions (in particular, their Assumption **A**), this family of programs (indexed by λ) leads to a simple characterization of the limit set of (perfect public) equilibrium payoffs as $\delta \to 1$. As our focus is the program itself, we shall not need Assumption **A** for what follows.

Denote by $k(\lambda) \in [-\infty, +\infty]$ the value of $\mathcal{P}(\lambda)$. HSTV prove that the feasible set of $\mathcal{P}(\lambda)$ is non-empty, so that $k(\lambda) > -\infty$, and that the value of $\mathcal{P}(\lambda)$ is finite, so that $k(\lambda) < +\infty$.

For a fixed Markov strategy (α_s), the feasible set is non-empty if and only if for all s, α_s is **admissible**, in the sense that, for all i, if there exists $\nu \in \Delta(A^i)$ such that, for all (t,y),

$$\sum_k \nu^k p(t,y|s,a^{ik},\alpha_s^{-i}) = p(t,y|s,\alpha_s),$$

then

$$\sum_k \nu^k r^i(s,a^{ik},\alpha_s^{-i}) \leq r^i(s,\alpha_s).$$

Indeed, it follows from Fan (1956) that there exists $x : S \times Y \to \mathbf{R}^{S\times I}$ such that, for each s, α_s is a Nash equilibrium of the game $\Gamma(s,x)$ if and only if α_s is admissible. Adding a constant to each $x_t(s,y)$ that is independent of (t,y), we may assume that v_s is independent of s. Finally, considering any i for which $\lambda^i \neq 0$, we may add (or subtract if $\lambda^i < 0$) a constant to $x_t^i(s,y)$, independent of s,t,y, so that the constraint **(ii)** is satisfied.

The purpose of this note is to prove that $\mathcal{P}(\lambda)$ is equivalent to (i.e. gives the same value as) $\tilde{\mathcal{P}}(\lambda)$ given by

$$\sup_{\alpha\in\Delta(A)^{|S|},\alpha_s \text{ admissible}} \min \sum_{s,i} \lambda^i \beta_s r^i \left(s, \hat{\alpha}_s^i, \alpha_s^{-i}\right),$$

where the minimum is over $(\hat{\alpha}_s^i)_{i,s}$, $\sum_k \hat{\alpha}_s^i \left(a^{ik}\right) = 1$, with $\lambda^i > 0 \Rightarrow (\alpha^i(a^{ik}) = 0 \Rightarrow \hat{\alpha}^i(a^{ik}) \leq 0, \alpha^i(a^{ik}) = 1 \Rightarrow \hat{\alpha}^i(a^{ik}) \geq 1)$, $\lambda^i < 0 \Rightarrow (\alpha^i(a^{ik}) = 0 \Rightarrow \hat{\alpha}^i(a^{ik}) \geq 0, \alpha^i(a^{ik}) = 1 \Rightarrow \hat{\alpha}^i(a^{ik}) \leq 1)$, and

$$\lambda^i \neq 0 \Rightarrow \hat{p}\left(t, y|s\right) := p\left(t, y|s, \hat{\alpha}_s^i, \alpha_s^{-i}\right) \geq 0,$$

as well as over $\beta_s \geq 0$, $\sum_s \beta_s = 1$ such that (β_s) is an invariant distribution of $\hat{p}\left(t \times Y|s\right)$. (If there are multiple invariant distributions, use the one that minimizes the objective function.)

Remark 1 *Cheng (2004)'s Theorem 5 corresponds to $\tilde{\mathcal{P}}(\lambda)$ with $|S| = 1$, where (β_s) collapses to the point mass. To our knowledge, Cheng is the first author to use duality to characterize the set of equilibrium payoffs in repeated games. Cheng works with the discounted program rather than the limit program. A related application of duality is Obara and Rahman (2010).*

Remark 2 *The sufficient conditions for a folk theorem to hold given by Fudenberg, Levine and Maskin (1994) follow immediately. As for those of Kandori and Matsushima (1998), note that their conditions can be stated in terms of convex cones: Adapting slightly their notation, let $Q^i(a) := \{p(\cdot|a^{-i}, a^{ik})|a^{ik} \in A^i \setminus \{a^i\}\}$ be the set of distributions over signals as player i's action varies all his actions but a^i. Let $C^i(a)$ denote the convex cone with vertex 0 spanned by $Q^i(a) - p(\cdot|a)$. Assumption **A2** of Kandori and Matsushima requires $C^i(a) \cap -C^j(a) = \{0\}$ and that 0 is not a conical combination of $Q^i(a) - p(\cdot|a)$, whereas Assumption **A3** requires $C^i(a) \cap C^j(a) = \{0\}$ for all $i \neq j$ and $a \in Ex(A)$ (the set of action profiles achieving some extreme point of the feasible payoff set). Note now that the restriction on $\hat{\alpha}$, when $\alpha = a$ is pure, is that $p(\cdot|\hat{\alpha}^i, a^{-i}) - p(\cdot|a) \in -C^i(a)$ whenever $\lambda^i > 0$, and $p(\cdot|\hat{\alpha}^i, a^{-i}) - p(\cdot|a) \in C^i(a)$ whenever $\lambda^i < 0$. Assumptions **A2** and **A3** then imply that $\hat{\alpha}^i = a^i$ for non-coordinate directions λ.*[1]

*As for stochastic games, HSTV's folk theorem follows immediately under **F1** and **F2**. It is also consistent with the average cost optimality equation (see Sennott, 1998), and Hoffman and Karp (1966). ($(\beta_s)_s$ is an invariant distribution after "deviations.")*

[1] For coordinate directions, admissibility suffices (cf. Kandori and Matsushima's Assumption **A1**).

2 Proof

Fix throughout some strategy (α_s) such that α_s is admissible for all s. We can rewrite $\mathcal{P}(\lambda)$ as

$$\max_{x,v} \lambda \cdot v$$

over x and v such that, for all s, i,

$$\sum_{t,y} p\left(t, y|s, \alpha_s\right) x_t^i\left(s, y\right) - v^i = -r^i\left(s, \alpha_s\right),$$

and, for all s, i, k,

$$\sum_{t,y} \left[p\left(t, y|s, a^{ik}, \alpha_s^{-i}\right) - p\left(t, y|s, \alpha_s\right)\right] x_t^i\left(s, y\right) \leq r^i\left(s, \alpha_s\right) - r^i\left(s, a^{ik}, \alpha_s^{-i}\right),$$

as well as, for all T, φ, ψ,

$$\lambda \cdot \sum_{s \in T} x_{\varphi(s)}\left(s, \psi\left(s\right)\right) \leq 0.$$

This is a linear program for (x, v). The first set of constraint ensures that α_s yields the same payoff v in all states, the second that playing α_s is a Nash equilibrium, and the third is the same constraint as **(ii)**. Because we assumed that α_s is admissible for all s, the feasible set is non-empty, and because the value of this program is bounded above by $k(\lambda)$, it is finite. We shall consider the dual of this linear program. It is

$$\min - \sum_{s,i} \gamma_s^i r^i\left(s, \alpha_s\right) + \sum_{s,i,k} \nu_s^{ik}\left(r^i\left(s, \alpha_s\right) - r^i\left(s, a^{ik}, \alpha_s^{-i}\right)\right)$$

over $\nu_s^{ik} \geq 0, \eta_{T\varphi\psi} \geq 0, \gamma_s^i \in \mathbf{R}$ such that, for all i, s, t, y,

$$p\left(t, y|s, \alpha_s\right) \gamma_s^i - \sum_k \left[p\left(t, y|s, \alpha_s\right) - p\left(t, y|s, a^{ik}, \alpha_s^{-i}\right)\right] \nu_s^{ik} + \lambda^i \sum_{T \supseteq \{s,t\}, \varphi(s)=t, \psi(s)=y} \eta_{T\varphi\psi} = 0,$$

where k runs through the actions of player i, and

$$\lambda^i = -\sum_s \gamma_s^i.$$

There is no loss in assuming $\lambda^i \neq 0$ for all i (we focus on the relevant subset of players otherwise). Define then $\beta_s^i := -\gamma_s^i/\lambda^i$ and $\xi_s^{ik} := \nu_s^{ik}/\lambda^i$. We get

$$\min \sum_{s,i} \lambda^i \left[\beta_s^i r^i(s, \alpha_s) + \sum_k \left(r^i(s, \alpha_s) - r^i\left(s, a^{ik}, \alpha_s^{-i}\right) \right) \xi_s^{ik} \right]$$

over $\eta_{T\varphi\psi} \geq 0, \xi_s^{ik}$ with $\xi_s^{ik} sgn(\lambda^i) \geq 0$, $\beta_s^i \in \mathbf{R}$ such that $\sum_s \beta_s^i = 1$ for all s, i, such that, for all (i, s, t, y),

$$\beta_s^i p(t, y|s, \alpha_s) + \sum_k \left(p(t, y|s, \alpha_s) - p\left(t, y|s, a^{ik}, \alpha_s^{-i}\right) \right) \xi_s^{ik} = \sum_{T \supseteq \{s,t\}, \varphi(s)=t, \psi(s)=y} \eta_{T\varphi\psi}.$$

Note that, taking the sum over (t, y),

$$\beta_s^i = \sum_{T \ni s, \varphi, \psi} \eta_{T\varphi\psi},$$

and so $\beta_s^i =: \beta_s$ is nonnegative and independent of i. Furthermore, by adding over s, we get that $\sum_{T,\varphi,\psi} |T| \eta_{T\varphi\psi} = 1$. Note also that, if $\beta_s = 0$ for some s, then $\sum_{T \ni s, \varphi, \psi} \eta_{T\varphi\psi} = 0$, and so, because $\eta_{T\varphi\psi} \geq 0$, it follows that

$$\sum_k \left(p(t, y|s, \alpha_s) - p\left(t, y|s, a^{ik}, \alpha_s^{-i}\right) \right) \xi_s^{ik} = 0;$$

Furthermore, because, given s and i, the variables ξ_s^{ik} all have the same sign independently of k, either $\sum_k \xi_s^{ik} \neq 0$, or $\xi_s^{ik} = 0$ for all k. Note that, in the former case, we can define the strategy $\tilde{\alpha}_s^i \in \Delta(A^i)$ by $\tilde{\alpha}_s^i(a^{ik}) = \xi_s^{ik} / \sum_{k'} \xi_s^{ik'}$, and admissibility then implies that the corresponding term in the objective function is nonnegative, and setting $\xi_s^{ik} = 0$ for all s would achieve a value at least as low. Hence, if $\beta_s = 0$ for some s, we can assume $\xi_s^{ik} = 0$ for all i, k, and the terms in the objective and the constraint that involve the state s vanish. Therefore, we might as well assume $\beta_s > 0$ for all s.

Define $\hat{\alpha}_s^i \in \mathbf{R}^{|A^i|}$ by, for all a^{ik},

$$\hat{\alpha}_s^i\left(a^{ik}\right) = \alpha_s^i\left(a^{ik}\right) + \frac{\alpha_s^i\left(a^{ik}\right)}{\beta_s} \sum_{k'} \xi_s^{ik'} - \frac{\xi_s^{ik}}{\beta_s}.$$

Note that $\sum_k \hat{\alpha}_s^i\left(a^{ik}\right) = 1$ for all s, i. We can rewrite our problem as

$$\min \sum_s \lambda^i \beta_s r^i\left(s, \hat{\alpha}_s^i, \alpha_s^{-i}\right),$$

over $(\hat{\alpha}_s^i)_{i,s}$, $\sum_k \hat{\alpha}_s^i\left(a^{ik}\right) = 1$, with $\lambda^i > 0 \Rightarrow (\alpha^i(a^{ik}) = 0 \Rightarrow \hat{\alpha}^i(a^{ik}) \leq 0, \alpha^i(a^{ik}) = 1 \Rightarrow \hat{\alpha}^i(a^{ik}) \geq 1)$, $\lambda^i < 0 \Rightarrow (\alpha^i(a^{ik}) = 0 \Rightarrow \hat{\alpha}^i(a^{ik}) \geq 0, \alpha^i(a^{ik}) = 1 \Rightarrow \hat{\alpha}^i(a^{ik}) \leq 1)$, as well as $\beta_s \geq 0$, $\sum_s \beta_s = 1$, and $\eta_{T\varphi\psi} \geq 0$, such that, for all (i, s, t, y),

$$\beta_s p\left(t, y | s, \hat{\alpha}_s^i, \alpha_s^{-i}\right) = \sum_{T \supseteq \{s,t\}, \varphi(s)=t, \psi(s)=y} \eta_{T\varphi\psi}. \tag{1}$$

Note that if $\beta_s > 0$, then it follows from (1) that $p\left(t, y | s, \hat{\alpha}_s^i, \alpha_s^{-i}\right)$ is nonnegative and independent of i. Also, if $\beta_s = 0$, we can assume $\hat{\alpha}_s^i = \alpha_s^i$ for all i without loss in the objective function. Thus in both cases, we can assume that $\hat{p}\left(t, y | s\right) := p\left(t, y | s, \hat{\alpha}_s^i, \alpha_s^{-i}\right) \geq 0$. Also note that (β_s) is an invariant distribution of transition function $\hat{p}\left(t \times Y | s\right)$. To see this, take the sum of (1) over s, y, and we have

$$\sum_s \beta_s \hat{p}\left(t \times Y | s\right) := \sum_{T \ni t, \varphi, \psi} \eta_{T\varphi\psi} = \beta_t.$$

Conversely, if (β_s) is an invariant distribution of $\hat{p}\left(t \times Y | s\right)$, then it follows from Lemma 1 of HSTV that there exists $\eta_{T\varphi\psi} \geq 0$ that satisfies (1).[2]

Thus we can rewrite our problem without using $\eta_{T\varphi\psi}$ as follows:

$$\min \sum_{s,i} \lambda^i \beta_s r^i\left(s, \hat{\alpha}_s^i, \alpha_s^{-i}\right),$$

over $(\hat{\alpha}_s^i)_{i,s}$, $\sum_k \hat{\alpha}_s^i\left(a^{ik}\right) = 1$, with $\lambda^i > 0 \Rightarrow (\alpha^i(a^{ik}) = 0 \Rightarrow \hat{\alpha}^i(a^{ik}) \leq 0, \alpha^i(a^{ik}) = 1 \Rightarrow \hat{\alpha}^i(a^{ik}) \geq 1)$, $\lambda^i < 0 \Rightarrow (\alpha^i(a^{ik}) = 0 \Rightarrow \hat{\alpha}^i(a^{ik}) \geq 0, \alpha^i(a^{ik}) = 1 \Rightarrow \hat{\alpha}^i(a^{ik}) \leq 1)$, and

$$\lambda_i \neq 0 \Rightarrow \hat{p}\left(t, y | s\right) := p\left(t, y | s, \hat{\alpha}_s^i, \alpha_s^{-i}\right) \geq 0$$

as well as $\beta_s \geq 0$, $\sum_s \beta_s = 1$ such that (β_s) is an invariant distribution of $\hat{p}\left(t \times Y | s\right)$. (If there are multiple invariant distributions, use the one that minimizes the objective function.) Taking the supremum over admissible $(\alpha_s)_s$, this gives us precisely $\tilde{\mathcal{P}}(\lambda)$.

[2] Use the indicator function of (s, t) for $(x_t(s))$ in the notation of Lemma 1. Note that one can easily generalize Lemma 1 to cases without irreducibility.

References

[1] Abreu, D., D. Pearce, and E. Stacchetti (1990). "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, **58**, 1041–1063.

[2] Cheng, H. (2004). "Analyzing Repeated Games with Imperfect Monitoring: A New Approach," *mimeo*, University of Southern California.

[3] Fan, K. (1956). "Systems of Linear Inequalities," in *Linear Inequalities and Related Systems*, H. W. Kuhn and A. W. Tucker, Ed., Paper 5. Annals of Mathematics Studies, Vol. 38. Princeton: Princeton Univ. Press.

[4] Fudenberg, D. and D. Levine (1994). "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, **62**, 103–135.

[5] Fudenberg, D., D. Levine, and E. Maskin (1994). "The Folk Theorem with Imperfect Public Information," *Econometrica*, **62**, 997–1040.

[6] Hoffman, A. J. and R. M. Karp (1966). "On Nonterminating Stochastic Games," *Management Science*, **12**, 359–370.

[7] Hörner, J., Sugaya, T., Takahashi S. and N. Vieille (2011). "Recursive Methods in Discounted Stochastic Games: An Algorithm for $\delta \to 1$ and a Folk Theorem," *Econometrica*, **79**, 1277–1318.

[8] Kandori, M., and H. Matsushima (1998). "Private observations, communication and collusion," *Econometrica*, **66**, 627–652.

[9] Rahman, D. and I. Obara (2010). "Mediated Partnerships," *Econometrica*, **78**, 285–308.

[10] Sennott, L. I. (1998). *Stochastic Dynamic Programming and the Control of Queuing Systems*, Wily Interscience, New York.