Institute for Empirical Research in Economics
University of Zurich

Working Paper Series
ISSN 1424-0459

Working Paper No. 215

Learning in Repeated Games without Repeating the Game

Patrick Leoni

December 2004

LEARNING IN REPEATED GAMES WITHOUT REPEATING THE GAME

PATRICK LEONI

Abstract. This paper extends the convergence result on Bayesian learning in Kalai and Lehrer (1993a, 1993b) to a class of games where players have a payoff function continuous for the product topology. Provided that 1) every player maximizes her expected payoff against her own beliefs, 2) every player updates her beliefs in a Bayesian manner, and 3) prior beliefs other players' strategies have a grain of truth, we show that after some finite time the equilibrium outcome of the above game is arbitrarily close to a Nash equilibrium. Those assumptions are shown to be tight.

1. Introduction

In their seminal works, Kalai and Lehrer (1993a, 1993b) give a set of sufficient conditions ensuring that Bayesian players, engaged in repeated interactions and learning about others' actions, end up behaving according to the prescription of a Nash equilibrium. The authors consider a class of games where every player has an utility function in the form of expected discounted sum of one-period payoffs. Every player has belief in the form of a probability distribution over opponents' strategies, and maximizes her payoff function against her belief. Provided that every player's belief has a grain of truth, Kalai and Lehrer show that the players' behaviors become arbitrarily close to those described in an almost-Nash equilibrium. Critical to the result of Kalai and Lehrer is that players must play the same game over time, and that their payoff functions are fully specified by discounting constant over time.

In this paper, we extend the above result to the case where players have payoff functions defined over the set of infinite play paths, with the additional assumption that payoff functions are continuous for the product topology. The class of games that we consider thus includes games where players may play different one-period games over time, or have history-dependent discount factors, or even do not play any one-period games but rather get their payoff at the end of every infinite history.

Even though we derive the same result as in Kalai and Lehrer for this class of games, the technique used here requires a significantly different approach. Indeed, the proof in Kalai and Lehrer (1993a, 1993b) relies on constant discounting and repetition of the same game to give an uniform upper-bound to the players' payoffs after some time. This allows to carry out the approximation of resulting

I would like to thank G. Carmona, C. Ewerhart and A. Rustichini for insightful conversations. This extends to the seminar participants at the University of Maastricht and the University of Minnesota. All remaining errors are mine.

play paths by a Nash equilibrium in a truncated game when beliefs are accurate enough through learning,[1] and to extend the approximation by an almost-Nash equilibrium to the infinite game by adding the control over subsequent payoff with the above upper-bound.

This approach does not apply when, for instance, players have payoffs defined over infinite histories. Still, continuity for the product topology used here allows for a somewhat similar control over future payoffs. We use this property to approximate resulting plays with a Nash equilibrium for a game with a finite number of histories derived from the original game (a somewhat equivalent notion of truncated games for our setting). This approximation is again ensured by continuity for the product topology for accurate enough beliefs, and global approximation by an almost-Nash equilibrium for the whole game is also a consequence of this assumption. The grain of truth together with Bayesian learning ensures convergence to accurate beliefs over time by a direct application of Blackwell-Dubins Theorem (see Blackwell and Dubins, 1962). Also, our simple continuity argument allows for a coarser topology of convergence towards Nash equilibrium than that in Kalai and Lehrer (1993a, 1993b).

We next go into more details to describe our result. We assume that players are *subjectively rational*; i.e., players have subjective beliefs about others' behaviors in the form of prior beliefs formed before the first period. Those prior beliefs are updated in every period in a Bayesian manner, according to available information. Every player is also assumed to maximize her expected intertemporal payoff against her subjective belief about other players' strategies.

Within this framework, we show that the actions taken by the players will become, for almost every infinite history, realization-equivalent to the actions taken in a Nash ε-equilibrium, for ε arbitrarily small. Moreover, this property occurs in finite time.

However, the above result relies on two important assumptions on the beliefs of the players, also used in Kalai and Lehrer (1993a, 1993b). The first critical assumption is that the prior beliefs of every player about others' actions have a *grain of truth*; i.e., the beliefs of every player assigns a strictly positive probability to every event that can occur with strictly positive probability during the game.[2] The second critical assumption is that every player believes that other players choose their actions independently of each others. In Section 5, we show that convergence fails when any of the last two assumptions is relaxed.

[1] Accuracy of beliefs is guaranteed by the combination of the grain of truth and Bayesian learning, as shown in Blackwell and Dubins (1962).

[2] In other words, every player assigns a strictly positive probability to every sequence of actions that can be chosen with strictly positive probability by the players.

We extend the concept of *subjective equilibrium*[3] as well as the concept of *self-confirming equilibrium*[4] to our class of games. We show that, in finite time and for almost every infinite history, players' behaviors become identical to behaviors described in a subjective equilibrium. Given the properties of subjective equilibria discussed in the above references, this last result provides a second decision-theoretic foundation for the concept of Nash equilibrium. We then show that any subjective equilibrium is realization-equivalent to an almost-Nash equilibrium for accurate enough beliefs. The proof of this last statement uses the continuity of the utility functions to ensure convergence of payoffs as beliefs become correct. Finally, eventual correctness of beliefs follows from the grain of truth assumption, together with Blackwell-Dubins' Theorem (see Blackwell and Dubins, 1962).

Interestingly enough, our result also extends to the case of learning about nature. Since nature can be reinterpreted as an additional player maximizing a constant utility function, learning about nature' choices after any finite history as well as others' behaviors is a direct consequence of our work.

The paper is organized as follows. In Section 2 we formally describe the class of games and the equilibrium concepts; in Section 3 we give the main result; Section 4 contains some intermediary results, and finally all the technical proofs are in the Appendix.

2. The model

2.1. **The game.** The model and some assumptions needed to obtain the main result of the paper are now described.

Time is discrete and continues forever. A period is denoted by the letter t ($t = 1, 2, 3, ...$). There are n players ($n \geq 1$), who plays forever. In every period t ($t \in \mathbb{N}$), every player i has a finite set of actions $\hat{\Sigma}_i^t$. Let $\hat{\Sigma}^t$ be defined as $\prod_{i=1}^{n} \hat{\Sigma}_i^t$, the set of *action combinations* available in this period. Let also $\Sigma^t = \prod_{p=1}^{t} \hat{\Sigma}^p$ denote the set of histories of length t, with t possibly infinite,[5] and let $\overline{H}$ be the set of all finite histories; i.e., the set $\overline{H} = \bigcup_{t \in \mathbb{N}} \Sigma^t$.

For every $s_t \in \Sigma^t$ ($t \in \mathbb{N}$), a cylinder with base s_t is defined to be the set $C(s_t) = \{s \in \Sigma^\infty | s = (s_t, ...)\}$ of all infinite histories whose t initial elements coincide with s_t. We define the set Γ_t ($t \in \mathbb{N}$) to be the $\sigma-$algebra that consists of all finite unions of cylinders with base on Σ^t; and Γ_0 is

[3]As defined in Kalai and Lehrer (1993a), see also Battigalli et al. (1992) for a history and a discussion of this concept.

[4]See Fudenberg and Levine (1993a).

[5]The set Σ^0 is defined to be the singleton consisting of the null history.

defined to be the trivial $\sigma-$algebra. The sequence $(\Gamma_t)_{t\in\mathbb{N}}$ generates a filtration, and we define Γ to be the $\sigma-$algebra generated by $\underset{t\in\mathbb{N}}{\cup}\Gamma_t$.

A *(behavioral) strategy* for every player assigns to every possible finite history a possibly randomized action. We represent a (behavioral) strategy for player i $(i = 1, ..., n)$ as a function

$$f : \overline{H} \longrightarrow \bigcup_{t\geq 0} \Delta(\hat{\Sigma}_i^t),$$

where $\Delta(\hat{\Sigma}_i^t)$ is the set of all probability distributions over $\hat{\Sigma}_i^t$ in every period t.

The game is played with *perfect monitoring*; i.e., the players know all realized past action combinations actually played.

2.2. **Realized play paths.** The concept of *infinite play path* is now defined.[6] This notion of infinite (or realized) play path represents the actual actions chosen by the players over time, in the following sense.

Consider a $n-$vector of behavioral strategies $f = (f_i)_{i=1,...,n}$. The null history h_0 leads to the realized action combination z^1 in the support of $f(h_0)$. Defined recursively, in period $t+1$, the players will choose the randomizations $f(z^t)$, which will result in the action combination z^{t+1}. The vector $(z^1, z^2, z^3, ...)$ is called the *realized play path.* A realized play path, finite or infinite, will be denoted by the letter z.

Denote by A_f the support of f; that is, the set of infinite play paths whose every finite truncation is assigned strictly positive probability by f.

2.3. **Beliefs.** The beliefs of the players about others' strategies are now formally described. Every player is assumed to have subjective prior beliefs about both other players' strategies. Those prior beliefs are formed before the first period of the game, and they will be updated in every subsequent period in a Bayesian manner, according to available information.

Formally, the beliefs of player i $(i = 1, ..., n)$ regarding the strategies of the other players are represented by a $n-$vector of strategies $f^i = (f^i_j)_{j=1,...,n}$. The belief f^i_j represents the belief of player i about player j' s strategy. Moreover, player i knows her own strategy; i.e., $f^i_i = f_i$ for every i.

We consider the following probabilistic representation of beliefs. We associate to a $n-$vector of strategies f a unique probability measure μ_f on the set on infinite play paths, as follows.

First, the measure μ_f is defined inductively on the set of finite play paths, and then uniquely extended to the set of infinite play paths. Define μ_f to be 1 for the null history. Consider now a

[6] What follows is described in Kalai and Lehrer (1993b).

finite history $h \in \overline{H}$ of length t, whose corresponding realized play path is given by the vector z, and a vector of actions $a \in \hat{\Sigma}^{t+1}$. The value $\mu_f(z, a)$ is inductively defined to be

$$\mu_f(z) \prod_i f_i(h)(a_i).$$

So defined, we now uniquely extend this measure to Σ^∞. Any finite history $h \in \overline{H}$, whose corresponding play path is given by z, is now consider as being the cylinder $C(z)$ (which is by definition the set of infinite paths whose initial segment is q).

Define the probability measure $\widetilde{\mu}_f$ as $\widetilde{\mu}_f(C(q)) \equiv \mu_f(h)$ for every such cylinders, and consider its unique extension to (Σ^∞, Γ) given by Caratheodory's Theorem. The extension of the probability measure $\widetilde{\mu}_f$ to (Σ^∞, Γ) is the unique extension of μ_f on (Σ^∞, Γ).[7]

The above representation implicitly requires that every player believes that other players choose their actions independently of each others. The reader is referred to Kalai and Lehrer (1993a) for a counterexample showing that none of the current results hold without this last assumption, and for a discussion of those assumptions on beliefs and their behavioral implications.

For sake of notational convenience, we shall denote by the same symbol f^i the prior belief others' strategies of player i and her updated beliefs obtained by iterated applications of Bayes' formula.

2.4. **Payoffs.** The intertemporal payoff functions of the players are now described. Every player has the utility function over the set of infinite histories

$$u^i : \Sigma^\infty \longrightarrow \mathbb{R}.$$

We assume that, for every player i,

1) the function u^i is continuous with respect to the product topology on Σ^∞, and

2) u^i is uniformly bounded above and below.

For any $(n-)$vector of strategies f, player i $(i = 1, ..., n)$ receives the expected payoff

$$U^i(f) = E^f \left[u^i(z) \right],$$

where E^f is the expected value with respect to the probability measure μ_f induced by the strategy f. Moreover, every player is assumed to maximize the above expression, namely her (subjective) expected payoff against her subjective belief about other players' strategies.

The above specification of payoffs encompasses the case treated in Kalai and Lehrer (1993a, 1993b) and Sandroni (1998), where the payoff over infinite histories takes the form of expected discounted sum of one-period payoff.

[7]See for instance Kalai and Lehrer (1993a) for more details.

2.5. **Equilibrium concepts.** This section is devoted to defining the solution concepts that will be used throughout. First, the concept of best-response against others' strategies is defined. Pick any player i $(i = 1, ..., n)$, and consider a $(n-1)$-vector strategies f_{-i} (this last vector can represent either beliefs about others' strategies, or actual strategies). For every $\alpha \geq 0$, a strategy f_i is a $\alpha-$*best response to* f_{-i} if

$$U^i(g, f_{-i}) - U^i(f_i, f_{-i}) \leq \alpha,$$

for every other strategy g available to player i.

Fix now $\alpha > 0$. A *Nash* $\alpha-$*equilibrium* is a $n-$vector f of strategies such that f_i is an $\alpha-$best response to f_{-i} for every i $(i = 1, ..., n)$. In particular, in any (almost) Nash equilibrium, beliefs about others' strategies are exact.

The next notion allows to specify a concept of closeness, in a probabilistic sense, between two vectors of strategies. Define first, for ψ and ψ' probability measures on the same probabilistic space $(\Omega, \mathcal{F})$, the *sup-norm* to be

$$\left\|\psi - \psi'\right\|_\infty = \sup_{A \in \mathcal{F}} \left|\psi(A) - \psi'(A)\right|.$$

Definition 1. Fix $\varepsilon > 0$. The strategy profile f plays $\varepsilon-$like the strategy profile g if

$$\left\|\mu_f - \mu_g\right\|_\infty < \varepsilon.$$

The concept of "playing $\varepsilon-$like" for two given strategies measures how distant those strategies are from each other in a probabilistic sense. It is preferable to approach this issue from a probabilistic standpoint, as explained in details in Kalai and Lehrer (1993a, 1993b), even though the authors use a different concept.

With the above definitions, it is now possible to introduce the concept of *subjective equilibrium* (up to some constants), which generalizes to our setting the concept of subjective equilibrium introduced in Kalai and Lehrer (1993a, 1993b), and the concept of self-confirming equilibrium introduced in Fudenberg and Levine (1993a, 1993b). This notion will play an important role in the proof of the main result of this paper.

Definition 2. Fix $\varepsilon > 0$. A $\varepsilon-$ subjective equilibrium is a matrix of beliefs $(f^j)_{j=1,\ldots,n}$, satisfying for every i $(i = 1, ..., n)$:

i) the strategy f^i_i is a best-response to f^i_{-i}, and

ii) the strategy profile $(f^j_j)_{j=1,\ldots,n}$ plays $\varepsilon-$like f^i.

In words, in any subjective equilibrium, the following requirements hold: *i*) every player maximizes her intertemporal utility function against her beliefs about others' strategies , and *ii*) the beliefs about others' strategies are realization-equivalent (up to ε) to actual plays.

3. The main result

In this section, the main result of the paper is stated and discussed. That is, the set of sufficient conditions leading to convergence toward Nash equilibria in finite time is given. We first introduce a definition, which captures the concept of *induced strategy* resulting from a given strategy after a particular finite history.

Definition 3. Consider a $n-$vector of strategies f, a period $t \in \mathbb{N}$ and a finite history $h \in \Sigma^t$. The induced strategy f_h is defined as

$$f_h(h') = f(hh') \text{ for any } h' \in H^r \ (r \in \mathbb{N}),$$

where hh' is the history of length $t+r$ resulting from the concatenation of the history h (first) and (followed by) the history h'.

For any p-vector of individual strategies $\widetilde{f} = (f_1, ..., f_p)$ (with $1 \leq p \leq n+1$), the induced $p-$vector of strategies $\widetilde{f}_h$ is defined as

$$\widetilde{f}_h = \left((f_1)_h, ..., (f_p)_h \right).$$

Before stating the main result of this paper, a notion in Measure Theory is first defined. Consider two measures λ and $\widetilde{\lambda}$ on the same measurable space $(\Omega, \mathcal{P})$. The measure λ is said to be *absolutely continuous* with respect to $\widetilde{\lambda}$, denoted by $\lambda \ll \widetilde{\lambda}$, if for every $A \in \mathcal{P}$ such that $\lambda(A) > 0$ it is true that $\widetilde{\lambda}(A) > 0$.

Finally, for any realized play path z and time $t > 0$, denote by $z(t)$ the truncation of z to its t first elements (thus $z(t) \in \Sigma^t$).

Theorem 1. Consider a $n-$vector of strategies f representing actual plays, and for every player j the belief f^j with $f^j_j = f_j$ such that, for every i,

i) the strategy f_i is a best-response to f^i_{-i},

ii) the beliefs are such that $\mu_f \ll \mu_{f^i}$, and

iii) player i updates her beliefs in a Bayesian manner.

Fix now any arbitrary $\alpha > 0$. For μ_f-almost every path z, there exists a time T such that, for every $t \geq T$, there exists a strategy profile $\overline{f}$ such that

1) $\overline{f}$ is a Nash $\alpha-$equilibrium, and

2) $\overline{f}$ plays 0-like f.

The above theorem says that, if 1) players maximize their intertemporal utility functions against their own beliefs, and if 2) beliefs are updated in a Bayesian manner, as long as the independence requirement is satisfied and the grain of truth holds, actual plays are realization-equivalent to an almost Nash equilibrium in finite time.

One of the key to proving the above result is that, when Assumptions *i)-iii)* are satisfied, along the realized play path actions satisfy the properties of a (almost-)subjective equilibrium in finite time. Since also any (almost) Nash equilibrium is an almost stochastic subjective equilibrium, and since also any (almost) Nash equilibrium trivially satisfies Assumptions *i)-iii)* above, Theorem 1 implicitly establishes some form of equivalence between those three different concepts.

The assumptions used in Theorem 1 are tight. For counterexamples when any of those assumptions is violated, the reader is referred to Kalai and Lehrer (1993a).

In terms of possible extensions to Theorem 1, it is conjectured that Bayesian learning is not the only (non-trivial) learning process for which the above result holds. The characterization of all learning processes for which convergence toward a Nash equilibrium obtains is an open problem.

The proof of Theorem 1 is given in the next section.

4. Proof and intermediary results

This section is devoted to giving the main line of the proof of Theorem 1. Since the proof is technical, only the main intermediary results are presented here.

The strategy of the proof of Theorem 1 goes as follows. First, it is shown that beliefs and strategies satisfying Assumptions *i)-iii)* become, in finite time, identical to an almost subjective equilibrium. Second, almost subjective stochastic equilibria are shown to be realization-equivalent to an almost Nash equilibrium when beliefs are accurate enough. Finally, it is shown that beliefs becomes accurate enough, in finite time. Overall, this leads to the approximation of initial strategies and beliefs by an almost Nash equilibrium, as in Theorem 1.

The first proposition makes the link between strategies and beliefs satisfying Assumptions *i)-iii)* in Theorem 1, and the concept of (almost-)subjective equilibrium. Its proof is given in the Appendix.

Proposition 2. Consider a $n-$vector of strategies f representing actual plays, and for every player j the belief f^j with $f^j_j = f_j$ such that, for every i,

i) the strategy f_i is a best-response to f^i_{-i},

ii) the beliefs are such that $\mu_f \ll \mu_{f^i}$, and

iii) player i updates her beliefs in a Bayesian manner.

For every $\varepsilon > 0$ and for μ_f-almost every play path z, there exists a time T such that, for every $t \geq T$, the strategy profile $(f^i_{z(t)})_{i=1,\ldots n}$ is a $\varepsilon-$subjective equilibrium for the repeated game starting after $z(t)$.

The above result implies that, as long as Assumptions *i)-iii)* hold, actual plays and beliefs about others' strategies along almost every path will become, in finite time, a (almost-) subjective equilibrium.

The next proposition makes the link between (almost-) subjective equilibrium and (almost-) Nash equilibrium. Its proof is also given in the Appendix.

Proposition 3. For every $\alpha > 0$, there exists $\overline{\varepsilon} > 0$ such that for every $\varepsilon < \overline{\varepsilon}$, if $\left(f^i\right)_{i=1,\ldots,n}$ is a $\varepsilon-$subjective equilibrium, then there exists a strategy profile $\overline{f}$ such that

1) $\overline{f}$ is a Nash $\alpha-$equilibrium, and

2) $\overline{f}$ plays 0-like f.

The above result mainly states that, provided that beliefs are accurate enough, every subjective equilibrium is an (almost-) Nash equilibrium. Given the conclusion of Proposition 3, and in order to prove the main result, it is enough to ensure that beliefs become arbitrarily correct.

Arbitrary accuracy of beliefs follows from the next proposition, which is the well-known and important result proved by Blackwell and Dubins (1962). It is a convergence result for conditional probabilities, stating that as information increases conditional probabilities of two different measures will convergence, as long as a requirement of absolute continuity is satisfied by those two measures.

Before stating the result, let $(\Omega, \mathcal{F})$ be a probabilistic space, and $(\mathcal{P}_t)_{t\in\mathbb{N}}$ be an increasing sequence of countable partitions of Ω, also called filter. This sequence of partitions represents the information available to an agent in any given period. For any $w \in \Omega$ and any period t, let $P_t(w)$ be the unique set in $\mathcal{P}_t$ such that $w \in P_t(w)$. Its proof can be found in Blackwell and Dubins (1962).

Theorem 4. (Blackwell-Dubins)

Consider two σ-additive measures μ and $\widetilde{\mu}$ on $(\Omega, \mathcal{F})$ such that $\mu \ll \widetilde{\mu}$. For $\mu-$almost every $w \in \Omega$ and for every $\varepsilon > 0$, there exists a time T such that

$$|\mu(A|P_t(w)) - \widetilde{\mu}(A|P_t(w))| < \varepsilon$$

for every $A \in \mathcal{F}$ and for every $t \geq T$.

With all the above intermediary results, we next move to the proof of Theorem 1.

Proof of Theorem 1.

Fix the strategies f and beliefs $\left(f^i\right)_{i=1,\ldots,n}$ such that $f_i = f_i^i$ for every i, and satisfying Assumptions *i)-iii)* of Theorem 1. Fix any $\alpha > 0$, and consider also $\overline{\varepsilon}$ associated with α and $\left(f^i\right)_{i=1,\ldots,n}$ by Proposition 3 and any $\varepsilon < \overline{\varepsilon}$.

By Proposition 2, for μ_f-every path there exists a time t_0 after which $\left(f^i\right)_{i=1,\ldots,n}$ is a $\varepsilon-$subjective equilibrium. By Proposition 3, there exists also a Nash α-equilibrium for the repeated game starting after $z(t_0)$ that plays 0-like f. Thus, we have found a period t_0 such that the original strategies play 0-like a Nash $\alpha-$ equilibrium in the repeated game starting after $z(t_0)$. The proof is now complete.

5. Appendix

The Appendix is devoted to proving technical results left aside earlier in the paper.

5.1. **Proof of Proposition 2.** The proof of this result goes as follows.

Fix $\varepsilon > 0$. Consider now Γ_t to be the set of all cylinders up to time t ($t \in \mathbb{N}$) and their extensions to the infinitely repeated-game; that is each element of Γ_t is the set of infinite play paths with the same basis of actions up to time t. Clearly, each Γ_t is a partition of the set of infinite play paths, and the family $(\Gamma_t)_{t\in\mathbb{N}}$ is a filtration of this set.

Therefore, Theorem 4 applies to the probability measures μ_f and μ_{f^i} for every i and to the filtration $(\Gamma_t)_{t\in\mathbb{N}}$. It follows that for μ_f-almost every infinite play path z, there exists a time t such that for every $s \geq t$ it is true that

$$\left|\mu_f(.|\Gamma_s(z)) - \mu_{f^i}(.|\Gamma_s(z))\right| < \varepsilon$$

for every i. It follows that $f_{z(t)}$ plays ε-like $f^i_{z(t)}$ for every i.

Consider any realized play paths z described above and the corresponding time t. By the Law of Iterated Expectations, the actions chosen by players i after the history $z(t)$, with the belief $\mu_{f^i}(.|\Gamma_t(z))$ and taking as given behaviors outside of $z(t)$, are identical to the actions chosen after the history $z(t)$ in the first period with belief μ_{f^i}. Since f_i is a best-response to f^i for every i, this implies that $(f_i)_{z(t)}$ is best-response to $\left(f^i_{-i}\right)_{z(t)}$ for every i. The proof of Proposition 2 is now complete.

5.2. **Proof of Proposition 3.** We first start with a technical lemma, stating that when two measures become eventually similar for the sup-norm, the expectations of any continuous functions according to those measures also become eventually similar. For any complete metric space $(\Omega, d(.))$, we denote by Φ the $\sigma-$algebra generated by the open ball of the metric and by Υ the topology generated by the same open balls.

Lemma 5. Consider two positive and finite measures λ and $\widetilde{\lambda}$ defined on the measurable space (Ω, Φ). Let $u : \Omega \longrightarrow \mathbb{R}$ be a continuous function for the topology Υ, uniformly bounded above and below. It is true that

$$\left|E^{\lambda}(u) - E^{\widetilde{\lambda}}(u)\right| \leq \sup_{x\in\Omega} |u(x)| \,.\, \left\|\lambda - \widetilde{\lambda}\right\|_{\infty}.$$

Proof. Consider any such function u, any such measures λ and $\widetilde{\lambda}$. We have that

$$\begin{aligned}\left|E^{\lambda}(u)-E^{\widetilde{\lambda}}(u)\right| &= \left|\int_{\Omega}u.d\lambda-\int_{\Omega}u.d\widetilde{\lambda}\right| \\ &= \left|\int_{\Omega}u.d\left(\lambda-\widetilde{\lambda}\right)\right| \\ &\leq \sup_{x\in\Omega}|u(x)|\,.\left\|\lambda-\widetilde{\lambda}\right\|_{\infty}.\end{aligned}$$

The proof is complete. □

We next state another technical lemma, related to the notion of subjective equilibrium. For every i ($i=1,...,n$), define first $\overline{u}^i=\sup_{s\in\Sigma^{\infty}}\left|u^i(s)\right|$, and then $\overline{u}=\max_{i=1,\ldots,n}\overline{u}^i$.

Lemma 6. For every $\alpha>0$, there exists $\widetilde{\varepsilon}>0$ such that, if $\left(f^j\right)_{j=1,\ldots,n}$ is a $\widetilde{\varepsilon}-$subjective equilibrium, then for every i and for every behavioral strategy g such that $A_{g,f_{-i}}\subseteq A_f$, the following holds:

$$\left|U^i(g,f_{-i})-U^i(g,f^i_{-i})\right|\leq\alpha.$$

Proof. To prove the result, we first truncate the infinite repeated game to a finitely repeated game, show that the result holds within this truncated game, and then extends the result to the original framework.

Fix $\alpha>0$. Consider any strategy vector $\left(f^j\right)_{j=1,\ldots,n}$, any i and any behavioral strategy g such that $A_{g,f_{-i}}\subseteq A_f$. First, we have that for every i, the function u^i is continuous for the product topology. This implies that there exists a period t_0 such that the contribution of any strategy profile to the overall payoff of every player after period t_0 is no greater than $\frac{\alpha}{4}$.

We restrict our attention to the truncated game of length t_0 in the following way: we consider the original strategy profile up to period t_0, and leave payoff constant thereafter by extending the original strategy profile to a constant arbitrary strategy profile after t_0. By our previous remark, the difference in payoff between the original strategy profile and the newly formed one is no greater than $\frac{\alpha}{4}$ for every player.

Formally, for any behavioral strategy q, we denote by q^{t_0} the restriction of q to Σ^{t_0}, and by q^{-t_0} the restriction of q to $\left(\Sigma^{t_0}\right)^c$. To truncate strategies, fix also any (dummy) strategy profile d such that $d_j(h_t)=d_j(h_{t'})$ for every $h_t,h_{t'}$ and j. For any strategy profile $p=(p_1,...,p_{n+1})$, define now for every j the truncated strategy $\overline{p}_j=\left(p_j^{t_0},d_j^{-t_0}\right)$, and let $\overline{p}=(\overline{p}_1,...,\overline{p}_{n+1})$. Consider also the function

$$\widetilde{U}^i(p)=U^i\left(\overline{p}\right).$$

With this last function, only changes of individual strategy within the truncated game of length t_0 can affect the value of $\widetilde{U}^i$.

In a first step, we show that $\left|\widetilde{U}^i(g,f_{-i}) - \widetilde{U}^i(g,f^i_{-i})\right| \leq \frac{\alpha}{2}$ for any unilateral deviation g from player i in the support of his initial strategy.

By applying Lemma 5 applied to $\widetilde{U}^i$, we get for every i and g such that $A_{g,f_{-i}} \subseteq A_f$ that

$$\left|\widetilde{U}^i(g,f_{-i}) - \widetilde{U}^i(g,f^i_{-i})\right| \leq \overline{u}.\left\|\mu_{\overline{g},\overline{f}_{-i}} - \mu_{\overline{g},\overline{f}^i_{-i}}\right\|_{\Sigma^{t_0}}, \tag{1}$$

where $\|.\|_{\Sigma^{t_0}}$ is the sup-norm restricted to the $\sigma-$algebra generated by Σ^{t_0}.

We next analyze the right-hand side of (1). To simplify notations, we define for every $s \in H^{T_0}$ the function

$$\Phi^i(s) = \left|\begin{array}{c} \prod\limits_{1\leq t\leq t_0-1}\prod\limits_{j\neq i} f_j(s^1,...,s^t)(s_j^{t+1}) \\ -\prod\limits_{1\leq t\leq t_0-1}\prod\limits_{j\neq i} f^i_j(s^1,...,s^t)(s_j^{t+1}) \end{array}\right|.$$

For every history $s \in H^{T_0}$, we have by construction of the beliefs that

$$\left|\mu_{\overline{g},\overline{f}_{-i}}(s) - \mu_{\overline{g},\overline{f}^i_{-i}}(s)\right| = \left|\Phi^i(s). \prod_{1\leq t\leq t_0-1} g(s^1,...,s^t)(s_i^{t+1})\right|.$$

Further, for any given $\varepsilon > 0$, if $\left(f^j\right)_{j=1,\ldots,n}$ is a $\varepsilon-$subjective equilibrium, we have for every history $s \in H^{T_0}$ s that

$$\left|\Phi^i(s). \prod_{1\leq t\leq t_0-1} f_i(s^1,...,s^t)(s_i^{t+1}).\right| \leq \varepsilon.$$

Consider now the set of such histories assigned strictly positive probability by f_i, and denote it by F. The last inequality implies for every $s \in F$ that

$$\Phi^i(s) \leq \frac{\varepsilon}{\prod\limits_{1\leq t\leq t_0-1} f_i(s^1,...,s^t)(s_i^{t+1})}.$$

Moreover, since $A_{g,f_{-i}} \subseteq A_f$, for every history $s \in F$ it must be true that $g(s) > 0$.

We next use the above remark to find an uniform upper-bound to the right-hand side of (1). Define

$$\rho = \min_{s\in F}\left|\prod_{1\leq t\leq t_0-1} f_i(s^1,...,s^t)(s_i^{t+1})\right|,$$

which is strictly positive, and let c denote the (finite) cardinal of Σ^{t_0}.

For every set B of finite histories of length t_0, from the above we have that

$$\begin{aligned}\left|\mu_{\overline{g},\overline{f}_{-i}}(B)-\mu_{\overline{g},\overline{f}^i_{-i}}(B)\right| &= \sum_{s\in B\cap F}\left|\mu_{\overline{g},\overline{f}_{-i}}(s)-\mu_{\overline{g},\overline{f}^i_{-i}}(s)\right| \\ &\leq \sum_{s\in B\cap F}\left(\Phi^i(s).\prod_{1\leq t\leq t_0-1} g(s^1,...,s^t)(s_i^{t+1})\right) \\ &\leq \sum_{s\in B\cap F}\Phi^i(s) \\ &\leq c\frac{\varepsilon}{\rho}.\end{aligned}$$

Taking the maximum over such sets, we have that

$$\left\|\mu_{\overline{g},\overline{f}_{-i}}-\mu_{\overline{g},\overline{f}^i_{-i}}\right\|_{\Sigma^{t_0}}\leq c\frac{\varepsilon}{\rho}.$$

Setting $\widetilde{\varepsilon}=\frac{\alpha\rho}{2\overline{u}}$, and together with (1), the previous analysis implies that if $\left(f^j\right)_{j=1,...,n}$ is a $\widetilde{\varepsilon}-$subjective equilibrium then

$$\left|\widetilde{U}^i(g,f_{-i})-\widetilde{U}^i(g,f^i_{-i})\right|\leq\frac{\alpha}{2}.$$

Moreover, since the contribution of any strategy profile after period t_0 is no greater than $\frac{\alpha}{4}$, if $\left(f^j\right)_{j=1,...,n}$ is a $\widetilde{\varepsilon}-$subjective equilibrium then

$$\begin{aligned}\left|U^i(g,f_{-i})-U^i(g,f^i_{-i})\right| &\leq \left|\widetilde{U}^i(g,f_{-i})-\widetilde{U}^i(g,f^i_{-i})\right|+\frac{\alpha}{2} \\ &\leq \alpha.\end{aligned}$$

We have thus derived the desired inequality, and the proof is now complete. □

With the two previous lemma, we can now prove Proposition 3. The proof goes as follows. Fix $\alpha>0$, and consider any vector of beliefs $\left(f^j\right)_{j=1,...,n}$. We associate to $\left(f^j\right)_{i=1,...,n}$ the following strategy profile $\overline{f}$:

1) for every $h\in A_f$, define $\overline{f}_i(h)=f_i(h)$ for every i,

2) for every $h\notin A_f$, consider the shortest prefix of h, say $\overline{h}$, such that $\overline{h}\notin A_f$ and consider two cases:

i) if $\overline{h}$ corresponds to an unilateral deviation by player j from the support of her strategy, define $\overline{f}_i(h)=f^j_i(h)$ for every $j\neq i$,

ii) if $\overline{h}$ does not correspond to an unilateral deviation, define $\overline{f}_i(h)$ arbitrarily.

To prove Proposition 3, it is enough to show that there exists $\overline{\varepsilon}>0$ such that, if $\left(f^j\right)_{J=1,...,n}$ is a $\overline{\varepsilon}-$subjective equilibrium , then its associated strategy profile $\overline{f}$ is a Nash $\alpha-$equilibrium.

We first claim that there exists $\varepsilon^1 > 0$ such that, if $\left(f^j\right)_{j=1,\dots,n}$ is a ε^1-subjective equilibrium, then for every i,

$$\left|U^i(f^i) - U^i(f)\ \right| \leq \frac{\alpha}{2}. \tag{2}$$

Indeed, by Lemma 5, we have for every i that

$$\left|U^i(f^i) - U^i(f)\ \right| \leq \overline{u}. \left\|\mu_{f^i} - \mu_f\right\|_\infty .$$

Define $\varepsilon^1 = \frac{1}{\overline{u}}\frac{\alpha}{2}$. Then for every ε^1-subjective equilibrium $\left(f^j\right)_{j=1,\dots,n}$ and for every i, the inequality (2) holds.

We next use the previous claim to get our result. In a first step we first prove the property for every individual deviation in the support of f, and then we extend this result to any arbitrary individual deviation.

By Lemma 6, there exists $\varepsilon^2 > 0$ such that, if $\left(f^j\right)_{j=1,\dots,n}$ is a ε^2-subjective equilibrium, then for every i and for every behavioral strategy g such that $A_{g,f_{-i}} \subseteq A_f$, the following holds:

$$\left|U^i(g, f_{-i}) - U^i(g, f^i_{-i})\right| \leq \frac{\alpha}{2}. \tag{3}$$

Define $\overline{\varepsilon} = \min(\varepsilon_1, \varepsilon_2)$. Clearly, for every $\overline{\varepsilon}-$subjective equilibrium $\left(f^j\right)_{j=1,\dots,n}$, for every i and every g such that $A_{g,f_{-i}} \subseteq A_f$, the following holds:

$$\begin{aligned} U^i(g, \overline{f}_i) - U^i(\overline{f}) \quad &= \quad U^i(g, f_i) - U^i(f) \\ &\quad\quad +U^i(g, f_i) - U^i(f^i) \\ &\quad\quad +U^i(f^i) - U^i(f). \end{aligned}$$

Combining (2) and (3) into this last relation gives

$$U^i(g, f_{-i}) - U^i(f)\ \leq \alpha + U^i(g, f^i) - U^i(f^i).$$

Moreover, since f_i is best response to (f^i_{-i}) for player i, the above implies that

$$U^i(g, f_{-i}) - U^i(f)\ \leq \alpha.$$

Equivalently, in terms of strategy profile $\overline{f}$, for any i and g such that $A_{g,f_{-i}} \subseteq A_f$ we have just shown that

$$U^i(g, \overline{f}_{-i}) - U^i(\overline{f})\ \leq \alpha.$$

We now extend this result to any arbitrary behavioral strategy g.

Fix any player i, and any strategy g. Assume that $A_{g,f_{-i}}$ differs from A_f. This implies that there exists an history h such that $g(h)$ is not in the support of $f_i(h)$. By construction of $\overline{f}$, in the subgames starting at any such corresponding unilateral deviation by player i, all the other players play according to f^i_{-i}. Since f_i is best-response to f^i_{-i} in those subgames, player i can improve upon g by playing in any such subgame according to f_i, and leave behaviors on A_f unchanged. We are therefore in the previous case, and the result follows.

All together, we have shown that $\overline{f}$ is a Nash α-equilibrium. Moreover, since as shown above there is no incentive to deviate from the original play paths, the strategy profile $\overline{f}$ plays 0-like f. The proof is now complete.

References

1. Aumann, R., 1964. Mixed and behaviour strategies in infinite extensive games. In: Dresher, M., Shapley, L., Tucker, A. (Eds.), Advances in Game Theory, Annals of Mathematics Studies 52. Princeton Univ. Press, Princeton, pp. 627-650.
2. Battigalli, P., Gilli, M., Molinari, M.C., 1992. Learning and convergence to equilibrium in repeated strategic interactions: an introductory survey. Ricerche Economiche 46, pp. 335-377.
3. Blackwell, D., Dubins, L., 1962. Merging of opinions with increasing information. Annals of Mathematical statistics 38, pp. 882-886.
4. Fudenberg, D., Levine, D., 1993a. Self-confirming equilibrium. Econometrica 61, pp. 523-545.
5. Fudenberg, D., Levine, D., 1993b. Steady-state learning and self-confirming equilibrium. Econometrica 61, pp. 547-573.
6. Kalai, E., Lehrer, E., 1993a. Rational learning leads to Nash equilibrium. Econometrica 61, pp. 1019-1045.
7. Kalai, E., Lehrer, E., 1993b. Subjective equilibrium in repeated games. Econometrica 61, pp. 1231-1240.
8. Kuhn, H., 1953. Extensive games and the problem of information. In Kuhn, H., Tucker, A. (Eds.), Contributions to the Theory of Games, Annals of Mathematics Studies 28. Princeton Univ. Press, Princeton, pp. 193-216.
9. Shiryayev, A., 1984. Probability. Springer-Verlang, New-York.

University of Zurich, Institute for Empirical Research in Economics, Winterthurerstrasse 30, CH-8006 Zurich, Switzerland

E-mail address: `pleoni@iew.unizh.ch`