Approachability in Infinite Dimensional Spaces

by

Ehud Lehrer[1]
School of Mathematical Sciences
Sackler Faculty of Exact Sciences
Tel Aviv University, Ramat Aviv
Tel Aviv 69978, Israel
e-mail: lehrer@math.tau.ac.il

First version: April 1997
February 27, 2002

Abstract. The approachability theorem of Blackwell (1956b) is extended to infinite dimensional spaces. Two players play a sequential game whose payoffs are random variables. A set C of random variables is said to be approachable by player 1 if he has a strategy that ensures that the difference between the average payoff and its closest point in C, almost surely converges to zero. Necessary conditions for a set to be approachable are presented.

[1]I acknowledge Eilon Solan for his helpful comments.

1 Introduction

Blackwell (1956b) considered a two-player sequential game where the payoffs at each round are vectors in a finite dimensional space, rather than numbers. A pre-specified set in the (vector) payoff space C is said to be approachable by player 1 if he has a strategy that ensures that the difference between the partial average payoff and its closest point in C, almost surely converges with time to zero. In contrast with the finite dimensional payoff space, as in Blackwell (1956b), the payoff space considered here is infinite dimensional. We present sufficient conditions that ensure *almost surely* approachability.

One may consider a vector-payoff game as finitely many games played simultaneously. In each round a player takes an action which applies to all games played. If there are no transferable payoffs from one game to another, the payoffs of the games are considered as one vector. The objective of player 1 then is to bring the average vector payoff into a set C. Blackwell treated the case where in each round all games are played. Here we also examine the case where not all games are played all the time. In each round, depending on the history, a different set of games is played. In terms of vector-payoffs, some coordinates may be inactive in some rounds. The relevant average is therefore the sum of past payoffs divided by the number of times a coordinate was previously active. Thus, the sum of payoffs is divided by a number that may vary with the coordinate. This fact imposes a difficulty in that it does not allow use of the multi-linearity of the inner product.

The approachability theorem has been applied extensively since its inception. Blackwell (1956a) himself noted that Hannan's (1957) no-regret theorem can be proven by using the approachability theorem. Aumann and Maschler (1995) used it to show that the uninformed player in a repeated game with incomplete information can guarantee at least what is then proven to be the value. Recently the approachability theory gained a revival due to the influential work of Foster and Vohra (1997 and 1999) on calibration and its relation to correlated equilibrium. Hart and Mas-Colell (2000) demonstrated an interactive learning process that converges to correlated equilibrium. In Hart and Mas-Colell (2001) they used the idea behind the geometric principle of approachability to introduce a large family of adaptive learning procedures. Rustichini (1999) proved a no-regret theorem for a case of imperfect monitoring. Spinat (2000) showed that any minimal approachable set is a B-set, that is, a set which satisfies the condition of Blackwell's theorem.

In his original paper Blackwell (1956b) also introduced the notion of weak approachability. Vieille (1992) used differential games with a fixed duration to study weak approachability in finite dimensional spaces.

As for approachability in large spaces, Lehrer (2001a) used it to show that there exists a prediction scheme that passes a large set of checking rules a la David (1982). Sandroni at al. (2000) extended this result to the case where the checking rules are prediction-based, that is, when an inspector can use rules that are based on current forecasting rather than on historical ones only. Lehrer (2001b) showed the existence of a regret-free strategy against infinitely many performance criteria. Lehrer (2001c) introduced an infinite game where in each round player 1 chooses a digit and player 2 a distribution over digits. Player 1 wins if the sequence of digits he chose during the game is normal with respect to the measure induced by the sequence of distributions player 2 chose through the game. Lehrer (2001c) proved by the approachability theorem in large spaces that player 1 has a pure winning strategy in this game. This strategy is in particular a procedure by which one can construct an extended normal number with respect to any distribution.

In this paper we separate the geometric aspects of approachability from the strategic aspects. The geometric principles behind approachability are introduced first (Section 3) and then applied to random-variable-payoff games (in Section 5). Section 6 is devoted to demonstrating the relation between the law of large numbers and the idea of approachability.

2 Approachability in an Infinite Dimensional Space

Consider a sequential game where at each stage player i chooses an action from a measurable set S_i, $i = 1, 2$. Let $(s_1^n, s_2^n) \in S_1 \times S_2$ denote the pair of actions taken at time n. A history of length n is a sequence $(s_1^1, s_2^1, s_1^2, s_2^2, ..., s_1^n, s_2^n)$. Histories of length n will later be denoted as h^n. The set of all finite histories is $H = \cup_n (S_1 \times S_2)^n$. For any $h^s, h^n \in H$ we say that $h^s < h^n$ if h^s is a prefix of h^n. Denote $\mathcal{H} = (S_1 \times S_2)^\infty$. For a given $h^\infty \in \mathcal{H}$ we denote by h^n its n^{th} prefix.

Let $(\Omega, \mu, \mathcal{F})$ be a probability space and χ be a function from H to the set of random variables defined on $(\Omega, \mu, \mathcal{F})$ that takes only values in $\{0, 1\}$. Thus, for every $h \in H$, $\chi(h)$ is a random variable defined over $(\Omega, \mu, \mathcal{F})$ that attains only two values, 0 or 1. When $\chi(h)(\omega) = 1$, we say that after the history h, ω is *active* and otherwise, that ω is *inactive*. The function χ is called the *indicator*. The payoff at time n after the history $h^{n-1} \in H$ of length $n - 1$, is determined by the pair of actions chosen at that time, (s_1^n, s_2^n). This payoff, denoted by $Y_n(h^n)$ (where, $h^n = (h^{n-1}, s_1^n, s_2^n)$), is a random variable defined on the probability space $(\Omega, \mu, \mathcal{F})$. Moreover, this variable satisfies the condition that for almost every $\omega \in \Omega$, $Y_n(h^n)(\omega) = 0$ if $\chi(h^{n-1})(\omega) = 0$. In other words, on an inactive ω the payoff is 0 (i.e., $Y_n(h^n)(\omega) = 0$). For any n denote $\tilde{\chi}(h^{n-1}) = \sum_{h^s < h^{n-1}} \chi(h^s)$. Thus, $\tilde{\chi}(h^{n-1})$ counts the number of times along h^{n-1} that points in Ω were active. Set,[2], $\overline{Y}_n(h^n) = \frac{\sum_{h^s < h^n} Y_s(h^s)}{\tilde{\chi}(h^{n-1})}$. $\overline{Y}_n(h^n)(\omega)$ is the average payoff over the times

[2] Here and throughout the entire paper $\frac{0}{0}$ is referred to as 0 .

that ω was active.

Let C be a closed set in $L_2 = L_2(\Omega, \mu, \mathcal{F})$. For $f \in L_2$ denote by $\text{Proj}_C(f)$ a closest point in C to f.[3]

Remark. The indicator determines whether a point ω is active or inactive in stage n based only on previous actions taken by both players. In fact the set up can be a bit more general than that. All the results are applied also to the case where the indicator depends not only on previous actions but also on the actual stage-action of player 1.

Definition 1 *For a given $h^\infty \in \mathcal{H}$ we say that the sequence $\{Y_n(h^n)\}$* ap-*proaches C if $\overline{Y}_n(h^n) - \text{Proj}_C(\overline{Y}_n(h^n))$ converges μ-almost surely to zero whenever $\chi(h^n) = 1$ infinitely often (i.e., $\mu(\overline{Y}_n(h^n) - \text{Proj}_C(\overline{Y}_n(h^n)) \to 0 | \chi(h^n)(\omega) = 1$ infinitely often$) = 1$, if $\mu(\chi(h^n)(\omega) = 1$ infinitely often$) > 0$).*

A strategy of player i is a function σ_i from H to $\Delta(S_i)$, the set of distributions over S_i. Any pair of strategies (σ_1, σ_2) induces a distribution over $\mathcal{H}$, $\lambda_{(\sigma_1,\sigma_2)}$.

Definition 2 *The set C is* approachable by player 1 *if there is a strategy σ_1 such that for every σ_2 the following holds: $\{Y_n(h^n)\}$ approaches C for $\lambda_{(\sigma_1,\sigma_2)}$-almost every h^∞.*

Adapting Blackwell's technique (see also Mertens, Sorin and Zamir (1994) Chapter II Section 4) to the infinite dimensional case would guarantee approachability only in the norm (of the payoff space). The reader should not confuse the "almost surely" in the approachability statement of Blackwell and the approaching "in probability" of the average payoffs. Blackwell showed that the distance, with respect to the Euclidean norm, between the average payoff and the approached set converges along almost every play-path to zero. Thus, the "almost surely" statement refers to the space of

[3]In case C is convex there is exactly one such point.

play-paths, while the convergence of the average payoffs is in the Euclidean norm. When the game payoffs are random variables defined on some probability space, convergence in the L_2 norm implies only convergence of the average payoff variables *in probability* and not almost surely. Proving convergence of the average payoffs almost surely requires a different technique than that used by Blackwell.

3 The Geometric Principle of Approachability

3.1 Converging to the origin

The inner product in $L_2 = L_2(\Omega, \mu, \mathcal{F})$ will be denoted by $\langle \cdot, \cdot \rangle$ (i.e., $\langle f, g \rangle = \int fgd\mu$) and $\|f\|^2 = \langle f, f \rangle$.

Definition 3 a. *A sequence* $\{g_n\}$ *of functions in* L_2 *is* μ-almost surely bounded *if there is a real-valued function* B *on* Ω *(not necessarily in* L_2*) such that* $\max_n |g_n| \leq B$ *μ-almost surely.*

b. $\{g_n\}$ *is* L_2-bounded *if there is a function real-valued function* B *in* L_2 *such that* $\max_n |g_n| \leq B$ *μ-almost surely.*

The following proposition resemble Theorem 4.2 in Kuipers and Niederreiter (1974). For the sake of completeness I provide a proof which uses an idea from Kuipers and Niederreiter (1974).

Proposition 1 *Suppose that*

(a) $\{\widetilde{g}_n\}$ *is a μ-almost surely bounded sequence of random variables in* L_2*;*

(b) $\widetilde{f}_n = \frac{\widetilde{g}_1 + \ldots + \widetilde{g}_n}{n}$*; and*

(c) $\sum \frac{\|\widetilde{f}_n\|^2}{n}$ *converges.*

Then, $\widetilde{f}_n$ converges μ-almost surely to zero.

Proof:

Since $\sum \frac{\|\widetilde{f}_n\|^2}{n}$ converges, there is an unbounded sequence of increasing numbers $\{\beta_n\}$, all greater than 1, such that $\sum \frac{\beta_n\|\widetilde{f}_n\|^2}{n}$ also converges. Set $M_1 = 1$ and for every $n > 1$,

$$M_{n+1} = \left[\frac{\beta_{M_n}}{\beta_{M_n} - 1} M_n\right] + 1. \tag{3.1}$$

For every n let k_n be an integer in the interval $M_n < k \leq M_{n+1}$ that $\|\widetilde{f}_k\|^2$ attains its minimum, that is, $\|\widetilde{f}_{k_n}\|^2 \leq \|\widetilde{f}_k\|^2$ for every $M_n < k \leq M_{n+1}$. Thus,

$$\|\widetilde{f}_{k_n}\|^2 \leq \frac{1}{M_{n+1} - M_n} \sum_{k=M_n+1}^{M_{n+1}} \|\widetilde{f}_k\|^2 \leq \frac{M_{n+1}}{M_{n+1} - M_n} \sum_{k=M_n+1}^{M_{n+1}} \frac{\|\widetilde{f}_k\|^2}{k}. \tag{3.2}$$

Due to the fact that $\{\beta_n\}$ increases to infinity and to (3.1), $\frac{M_{n+1}}{M_{n+1}-M_n} < \beta_{M_n}$ for sufficiently large n. Furthermore, (3.2) implies

$$\|\widetilde{f}_{k_n}\|^2 \leq \sum_{k=M_n+1}^{M_{n+1}} \frac{\beta_k \|\widetilde{f}_k\|^2}{k}.$$

Thus, $\sum_{n=1}^{\infty} \|\widetilde{f}_{k_n}\|^2 = \sum_{n=1}^{\infty} \int (\widetilde{f}_{k_n})^2 d\mu < \infty$. By Fatou's lemma, $\int \sum_{n=1}^{\infty} (\widetilde{f}_{k_n})^2 d\mu < \infty$. Thus, $\sum_{n=1}^{\infty} (\widetilde{f}_{k_n})^2 < \infty$ μ-almost surely and $\widetilde{f}_{k_n} \to_n 0$ μ-almost surely. Due to (3.1), $\lim_n \frac{M_{n+1}}{M_n} = 1$ and thus $\lim_n \frac{k_{n+1}}{k_n} = 1$.

By (a) there is a function B such that $|\widetilde{g}_n| \leq B$ for every n. Therefore, if $M_n < k \leq M_{n+1}$, then (a) and (b) imply,

$$|\widetilde{f}_k| \leq \left|\frac{k_n \widetilde{f}_{k_n}}{k}\right| + \left|\frac{M_{n+1} - M_n}{M_n} B\right| \to_n 0 \quad \mu - \text{almost surely}. \tag{3.3}$$

This completes the proof. ∎

3.2 The principle of approachability: A simple version

Theorem 1 *Suppose that*

(a) $\{\tilde{X}_n\}_0^\infty$ *is a sequence of non-decreasing random variables that assume integer values,* $\tilde{X}_n - \tilde{X}_{n-1} \leq 1$, $\tilde{X}_n \to \infty$ μ*-a.s. and* $\tilde{X}_0 = 0$;

(b) $\{g_n\}$ *is an* L_2*-bounded sequence of random variables in* $L_2(\Omega, \mu, \mathcal{F})$ *that satisfies* $g_n = 0$ *whenever* $\tilde{X}_n - \tilde{X}_{n-1} = 0$;

(c) $f_n = \frac{\tilde{X}_{n-1} f_{n-1} + g_n}{\tilde{X}_n}$, *where* $f_1 = g_1$*; and*

(d) $\sum \langle \frac{f_{n-1}}{\tilde{X}_n}, g_n \rangle < \infty$.

Then, f_n *converges* μ*-almost surely to zero.*

Proof: Denote $A_n = \{\tilde{X}_n - \tilde{X}_{n-1} = 1\}$. Let $x_n = \mathbb{1}_{A_n} \frac{g_n - f_{n-1}}{\tilde{X}_n}$. Note that $f_n = \frac{\tilde{X}_{n-1} f_{n-1} + g_n}{\tilde{X}_n} = f_{n-1} + x_n$. Thus,

$$\|f_n\|^2 = \|f_{n-1}\|^2 + 2\langle f_{n-1}, x_n \rangle + \|x_n\|^2 =$$

$$\|f_{n-1}\|^2 + 2\langle f_{n-1}, \mathbb{1}_{A_n} \frac{g_n}{\tilde{X}_n} \rangle - 2\langle f_{n-1}, \mathbb{1}_{A_n} \frac{f_{n-1}}{\tilde{X}_n} \rangle + \|x_n\|^2.$$

By (b), $\mathbb{1}_{A_n} \frac{g_n}{\tilde{X}_n} = \frac{g_n}{\tilde{X}_n}$, and therefore,

$$\|f_n\|^2 = \|f_{n-1}\|^2 + 2\langle f_{n-1}, \frac{g_n}{\tilde{X}_n} \rangle - 2\langle f_{n-1}, \mathbb{1}_{A_n} \frac{f_{n-1}}{\tilde{X}_n} \rangle + \|x_n\|^2.$$

Continuing inductively we obtain,

$$(3.4) \quad \|f_n\|^2 = \|f_1\|^2 + \sum_{i=2}^{n} 2\langle f_{i-1}, \frac{g_i}{\tilde{X}_i} \rangle - \sum_{i=2}^{n} 2\langle f_{i-1}, \mathbb{1}_{A_i} \frac{f_{i-1}}{\tilde{X}_i} \rangle + \sum_{i=2}^{n} \|x_i\|^2.$$

Let B be the L_2 function that bounds g_n (see (b) of the theorem). Observe that,

(3.5)

$$\sum_{i=2}^{\infty}\|x_i\|^2 \leq \sum_{i=1}^{\infty}\|\mathbb{1}_{A_i}\frac{2B}{\tilde{\chi}_i}\|^2 = 4\sum_{i=1}^{\infty}\int_{A_i}\frac{\|B\|^2}{\tilde{\chi}_i^2}d\mu = 4\int\sum_{i=1}^{\infty}\frac{\|B\|^2}{i^2}d\mu < \infty.$$

Note that $\langle f_{i-1}, \mathbb{1}_{A_i}\frac{f_{i-1}}{\tilde{\chi}_i}\rangle$ is non-negative for every i. Condition (d), (3.4) and (3.5) imply that $\sum_{i=2}^{\infty} 2\langle f_{i-1}, \mathbb{1}_{A_i}\frac{f_{i-1}}{\tilde{\chi}_i}\rangle$ converges.

Define $j(n) = \min\{m;\ \tilde{\chi}_m \geq n\}$. $j(n)$ is the first stage m, where $\tilde{\chi}_m$ is equal to n. Due to (a), $j(n)$ is μ-almost surely well defined for every n. Set $\widetilde{g}_n = g_{j(n)}$ and define $\widetilde{f}_n = \frac{1}{n}\sum_{i=1}^{n}\widetilde{g}_n$. Note that the series $\sum_{i=2}^{\infty}\langle f_{i-1}, \mathbb{1}_{A_i}\frac{f_{i-1}}{\tilde{\chi}_i}\rangle$ is equal to $\sum_{n=1}^{\infty}\frac{\|\widetilde{f}_n\|^2}{n+1}$. Thus, the latter series converges. It implies that the series $\sum_{n=1}^{\infty}\frac{\|\widetilde{f}_n\|^2}{n}$ also converges. Therefore, (c) of Proposition 1 is satisfied.

The definition of $\widetilde{f}_n$ and the fact that the sequence $\widetilde{g}_n$ is μ-a.s. bounded (since $\{g_n\}$ is an L_2-bounded sequence) guarantee that (a) and (b) of Proposition 1 are also satisfied. Proposition 1 therefore implies that $\widetilde{f}_n$ converges μ-a.s. to zero. Thus, f_n converges to zero μ-almost surely, as desired. ■

A slight variation of Theorem 1 is the following corollary which will be useful later.

Corollary 1 *Suppose that*

(a) $\{\tilde{\chi}_n\}_0^{\infty}$ *is a sequence of non-decreasing random variables that assume integer values.* $\tilde{\chi}_n - \tilde{\chi}_{n-1} \leq 1$ *and* $\tilde{\chi}_0 = 0$*;*

(b) $\{g_n\}$ *is a* L_2*-bounded sequence of random variables such that* $\tilde{\chi}_n - \tilde{\chi}_{n-1} = 0$ *implies* $g_n = 0$*;*

(c) $f_n = \frac{\tilde{\chi}_{n-1}f_{n-1}+g_n}{\tilde{\chi}_n}$*; and*

(d) $\sum\langle\frac{f_{n-1}}{\tilde{\chi}_n}, g_n\rangle < \infty$*.*

Then, $\tilde{\chi}_n \to \infty$ implies f_n converges to zero μ-a.s. (i.e., the event where $\tilde{\chi}_n \to \infty$ and f_n does not converge to zero has probability zero).

Proof. One can redefine the sequence $\tilde{\chi}_n$ so that $\tilde{\chi}_n \to \infty$ μ-a.s. without changing the other random variables. Conditions (a)-(d) of Theorem 1 are therefore satisfied and Corollary 1 follows. ■

3.3 The principle of approachability: Converging to the origin with unbounded variables

In this section we deal with a special case where $\chi_n \equiv 1$. In this case we state a theorem analogous to Theorem 1 with unbounded random variables.

In what follows $\{g_n\}$ is a sequence of random variables and $f_n = \frac{g_1+...+g_n}{n}$.

Proposition 2 *Suppose that there exists $\delta > 0$ such that*

(a) $\|g_n\|^2 = O(n^{\frac{1}{3}-\delta})$, and

(b) $\|f_n\|^2 = O(n^{-\frac{2}{3}-\delta})$,

then, f_n converges μ-almost surely to zero.

Proof. Let $\beta_n = n^{\frac{2}{3}}$. Due to (b), $\sum \frac{\beta_n \|f_n\|^2}{n}$ converges. As in Proposition 1, set $M_1 = 1$ and for every $n > 1$, $M_{n+1} = \left[\frac{\beta_{M_n}}{\beta_{M_n}-1} M_n\right] + 1$. Recall k_n from the proof of Proposition 1 and the fact that $\frac{k_{n+1}-k_n}{k_n} \to_n 0$ μ-almost surely. Similarly to the proof of Proposition 1 we obtain $f_{k_n} \to_n 0$ μ-almost surely. To complete the proof it remains to show that $f_k \to_k 0$ μ-almost surely.

By (a), $\sum \frac{\|g_n\|^2}{n^{\frac{4}{3}}}$ converges. Therefore, by Fatou's lemma, $\int \sum \frac{(g_n)2}{n^{\frac{4}{3}}} d\mu$ converges, which implies that $\sum \frac{(g_n)2}{n^{\frac{4}{3}}}$ converges μ-almost surely. Thus,

$$\frac{g_n}{n^{\frac{2}{3}}} \to 0 \quad \mu - \text{almost surely.} \tag{3.6}$$

Let k satisfy $k_n < k \leq k_{n+1}$. Note that,

$$f_k = \frac{k_n}{k} f_{k_n} + \frac{k - k_n}{k} \frac{g_{k_n+1} + ... + g_k}{k - k_n}. \tag{3.7}$$

Since $M_n \leq k_n$ and $k_{n+1} \leq M_{n+2}$, there is a constant c such that

$$\frac{k-k_n}{k}\ \frac{g_{k_n+1}+...+g_k}{k-k_n} \leq \frac{M_{n+2}-M_n}{M_n}\ \frac{g_{k_n+1}+...+g_k}{k-k_n} \leq$$

$$\frac{c}{(M_n)^{\frac{2}{3}}}\ \frac{g_{k_n+1}+...+g_k}{k-k_n} \leq \frac{c(M_{n+2})^{\frac{2}{3}}}{(M_n)^{\frac{2}{3}}}\ \frac{1}{k-k_n}\ \left(\frac{g_{k_n+1}}{(M_{n+2})^{\frac{2}{3}}}+...+\frac{g_k}{(M_{n+2})^{\frac{2}{3}}}\right).$$

The first term tends to c, while the rest is an average of $k-k_n$ numbers that, by (3.6), go to zero μ-almost surely. Therefore, the right-hand term of (3.7) tends to zero and thus, $f_k \to 0$ μ-almost surely and the proof is complete. ■

Theorem 2 *Suppose that there exists $\delta > 0$ such that*

(a) $\{g_n\}$ is a sequence of random variables such that $\|g_n\|^2 = O(n^{\frac{1}{3}-\delta})$, and

(b) for every n, $\sum_{i=2}^{n}(i-1)\langle f_{i-1}, g_i\rangle = O(n^{\frac{4}{3}-\delta})$,

then, f_n converges μ-almost surely to zero.

Proof.

$$\|f_n\|^2 = \frac{(n-1)^2}{n^2}\|f_{n-1}\|^2 + 2\frac{(n-1)}{n^2}\langle f_{n-1}, g_n\rangle + \frac{1}{n^2}\|g_n\|^2.$$

Continuing inductively we obtain,

$$\|f_n\|^2 = \frac{1}{n^2}\|f_1\|^2 + \frac{2}{n^2}\sum_{i=2}^{n}(i-1)\langle f_{i-1}, g_i\rangle + \frac{1}{n^2}\sum_{i=2}^{n}\|g_i\|^2.$$

Due to (a) and (b),

$$\|f_n\|^2 = O(n^{-\frac{2}{3}-\delta}).$$

We conclude that the conditions of Proposition 2 hold, and therefore, f_n converges μ-almost surely to zero, as desired. ■

Corollary 2 *Suppose that*

(a) $\{g_n\}$ *is a sequence of random variables whose norm is uniformly bounded, and*

(b) $\langle f_{n-1}, g_n\rangle \leq 0$ *for every* n*,*

then, f_n *converges* μ*-almost surely to zero.*

Proof.

If (b) holds, then $\sum_{i=2}^{n}(i-1)\langle f_{i-1}, g_i\rangle \leq 0$ and therefore (b) of Theorem 2 holds. Since (a) of Theorem 2 is implied by (a), f_n converges μ-almost surely to zero as desired. ■

3.4 The principle of approachability: Converging to a closed set

In this section we describe the approachability principle as applied to a closed set C in L_2. Recall that $\mathrm{Proj}_C(f)$ is a closest point to f in C. Here also the discussion is confined to the case where $\chi_n \equiv 1$.

Theorem 3 *Suppose that* $\{g_n\}$ *is a sequence of random variables. Denote* $\overline{g}_n = \frac{g_1+\ldots+g_n}{n}$*. If there exists* $\delta > 0$ *such that*

(a) $\|g_n\|^2 = O(n^{\frac{1}{3}-\delta})$*, and*

(b) for every n*,* $\sum_{i=2}^{n}(i-1)\langle \overline{g}_{i-1} - \mathrm{Proj}_C(\overline{g}_{i-1}), g_i - \mathrm{Proj}_C(\overline{g}_{i-1})\rangle = O(n^{\frac{4}{3}-\delta})$*,*

then, $f_n = \overline{g}_n - \mathrm{Proj}_C(\overline{g}_n)$ *converges* μ*-almost surely to zero.*

Proof.

$$\|f_n\|^2 = \|\overline{g}_n - \text{Proj}_C(\overline{g}_n)\|^2 \leq \|\frac{(n-1)\overline{g}_{n-1}}{n} + \frac{g_n}{n} - \text{Proj}_C(\overline{g}_{n-1})\|^2 \leq$$
$$\Big\|\frac{n-1}{n}\Big(\overline{g}_{n-1} - \text{Proj}_C(\overline{g}_{n-1})\Big)\Big\|^2 +$$
$$2\Big\langle\frac{n-1}{n^2}\Big(\overline{g}_{n-1} - \text{Proj}_C(\overline{g}_{n-1})\Big), g_n - \text{Proj}_C(\overline{g}_{n-1})\Big\rangle + \|\frac{g_n - \text{Proj}_C(\overline{g}_{n-1})}{n}\|^2 =$$
$$(\frac{n-1}{n})^2\|f_{n-1}\|^2 +$$
$$2\frac{n-1}{n^2}\Big\langle\overline{g}_{n-1} - \text{Proj}_C(\overline{g}_{n-1}), g_n - \text{Proj}_C(\overline{g}_{n-1})\Big\rangle + \|\frac{g_n - \text{Proj}_C(\overline{g}_{n-1})}{n}\|^2.$$

Continuing inductively we obtain,

$$\|f_n\|^2 \leq \frac{1}{n^2}\|f_1\|^2 + \tag{3.8}$$
$$\frac{2}{n^2}\sum_{i=2}^{n}(i-1)\langle\overline{g}_{i-1} - \text{Proj}_C(\overline{g}_{i-1}), g_i - \text{Proj}_C(\overline{g}_{i-1})\rangle + \frac{1}{n^2}\sum_{i=2}^{n}\|g_i - \text{Proj}_C(\overline{g}_{i-1})\|^2.$$

Fix a point $c \in C$. Note that for every random variable x, $\|\text{Proj}_C(x)\| = \|x + \text{Proj}_C(x) - x\| \leq \|x\| + \|\text{Proj}_C(x) - x\|$ and $\|\text{Proj}_C(x) - x\| \leq \|c - x\| \leq \|c\| + \|x\|$. Thus, $\|\text{Proj}_C(x)\| \leq 2\|x\| + \|c\|$. Applying it to $x = \overline{g}_{i-1}$, (a) implies that $\|\text{Proj}_C(\overline{g}_{i-1})\|^2 = O((i-1)^{\frac{1}{3}-\delta})$, $i = 2, 3, \ldots$. Therefore

$$\frac{1}{n^2}\sum_{i=2}^{n}\|g_i - \text{Proj}_C(\overline{g}_{i-1})\|^2 = O(n^{-\frac{2}{3}-\delta}). \tag{3.9}$$

Due to (3.8), (3.9) and (b),

$$\|f_n\|^2 = O(n^{-\frac{2}{3}-\delta}). \tag{3.10}$$

In order to use Proposition 2, it remains to show that f_n is an average of random variables that satisfy (a) of that proposition.

Define $h_1 = g_1 - \text{Proj}_C(g_1)$ and $h_i = if_i - i\text{Proj}_C(f_i) - (i-1)f_{i-1} + (i-1)\text{Proj}_C(f_{i-1})$, $i = 2, 3, \ldots$. Obviously $f_n = \frac{h_1 + \ldots + h_n}{n}$. Due to (3.10),

$\|\text{Proj}_C(f_i)\|^2 = O(i^{-\frac{2}{3}-\delta})$. Thus, $\|h_i\| \leq i\|f_i\| + i\|\text{Proj}_C(f_i)\| + (i-1)\|f_{i-1}\| + (i-1)\|\text{Proj}_C(f_{i-1})\|$, which implies that $\|h_i\|^2 = O(i^{\frac{1}{3}-\delta})$ for every i.

From Proposition 2 we now infer that f_n converges μ-almost surely to zero and the proof is complete. ■

3.5 Convergence to a closed set: Unbounded random variables and $\chi_n \not\equiv 1$

Definition 4 *Let C be a closed set in $L_2 = L_2(\Omega, \mu, \mathcal{F})$. We say that C respects L_2-boundedness if*

***a**. whenever $\{r_n\}$ is bounded by an L_2 function so is the sequence $\{\text{Proj}_C(r_n)\}$;*

***b**. if for a sequence $\{r_n\}$ of L_2 functions there exists a function B such that $|r_n - r_{n-1}| < \frac{B}{n}$ for every n, then there exists a function B' such that $|\text{Proj}_C(r_n) - \text{Proj}_C(r_{n-1})| < \frac{B'}{n}$ for every n.*

Remarks.

a. Let f and g be two functions in L_2 and $W \in \mathcal{F}$. Then, the sets $C = \{h \in L_2\ ;\ \ g \leq h \leq f\}$, $C = \{h \in L_2\ ;\ \ h \leq f\}$ and $C = \{h \in L_2\ ;\ \ h\mathbb{1}_W \leq f\mathbb{1}_W\}$ respect L_2-boundedness, where $\mathbb{1}_W$ is the characteristic function of W. Thus, a point in L_2, the positive and the negative orthants of L_2 all respect L_2-boundedness.

b. If C is finite dimensional then C respects L_2-boundedness.

Theorem 4 *Suppose that C respects L_2-boundedness and that*

(a) $\{\tilde{\chi}_n\}_0^\infty$ is a sequence of non-decreasing random variables that assume integer values, $\tilde{\chi}_n - \tilde{\chi}_{n-1} \leq 1$, $\tilde{\chi}_n \to \infty$ μ-a.s. and $\tilde{\chi}_0 = 0$;

(b) $\{g_n\}$ is L_2-bounded and satisfies $g_n = 0$ whenever $\tilde{\chi}_n - \tilde{\chi}_{n-1} = 0$;

(c) $\overline{g}_n = \frac{\tilde{\chi}_{n-1}\overline{g}_{n-1}+g_n}{\tilde{\chi}_n}$, *where* $\overline{g}_1 = g_1$*; and*

(d) $\limsup_n \sum_{s\le n} \Big\langle \overline{g}_{s-1} - \mathrm{Proj}_C(\overline{g}_{s-1}), \chi_s \frac{g_s - \mathrm{Proj}_C(\overline{g}_{s-1})}{\tilde{\chi}_s} \Big\rangle < \infty$.

Then, $\overline{g}_n - \mathrm{Proj}_C(\overline{g}_n)$ *converges to zero* μ*-almost surely.*

Proof. Denote $\chi_n = \tilde{\chi}_n - \tilde{\chi}_{n-1}$.

$$\|\overline{g}_n - \mathrm{Proj}_C(\overline{g}_n)\|^2 \le \|\overline{g}_{n-1} + \chi_n\Big(\frac{g_n}{\tilde{\chi}_n} - \frac{\overline{g}_{n-1}}{\tilde{\chi}_n}\Big) - \mathrm{Proj}_C(\overline{g}_{n-1})\|^2 =$$

$$\|\overline{g}_{n-1} - \mathrm{Proj}_C(\overline{g}_{n-1})\|^2 + \|\chi_n \frac{g_n - \mathrm{Proj}_C(\overline{g}_{n-1})}{\tilde{\chi}_n}\|^2 +$$

$$\|\chi_n \frac{\mathrm{Proj}_C(\overline{g}_{n-1}) - \overline{g}_{n-1}}{\tilde{\chi}_n}\|^2 + 2\Big\langle \overline{g}_{n-1} - \mathrm{Proj}_C(\overline{g}_{n-1}), \chi_n \frac{g_n - \mathrm{Proj}_C(\overline{g}_{n-1})}{\tilde{\chi}_n}\Big\rangle +$$

$$2\Big\langle \overline{g}_{n-1} - \mathrm{Proj}_C(\overline{g}_{n-1}), \chi_n \frac{\mathrm{Proj}_C(\overline{g}_{n-1}) - \overline{g}_{n-1}}{\tilde{\chi}_n}\Big\rangle +$$

$$2\Big\langle \chi_n \frac{\mathrm{Proj}_C(\overline{g}_{n-1}) - \overline{g}_{n-1}}{\tilde{\chi}_n}, \frac{g_n - \mathrm{Proj}_C(\overline{g}_{n-1})}{\tilde{\chi}_n}\Big\rangle.$$

Continuing inductively we obtain,

$$\|\overline{g}_n - \mathrm{Proj}_C(\overline{g}_n)\|^2 \le \tag{3.11}$$

$$\|g_1 - \mathrm{Proj}_C(g_1)\|^2 + \sum_{s\le n} \|\chi_s \frac{g_s - \mathrm{Proj}_C(\overline{g}_{s-1})}{\tilde{\chi}_s}\|^2 +$$

$$\sum_{s\le n} \|\chi_s \frac{\mathrm{Proj}_C(\overline{g}_{s-1}) - \overline{g}_{s-1}}{\tilde{\chi}_s}\|^2 + 2\sum_{s\le n} \Big\langle \overline{g}_{s-1} - \mathrm{Proj}_C(\overline{g}_{s-1}), \chi_s \frac{g_s - \mathrm{Proj}_C(\overline{g}_{s-1})}{\tilde{\chi}_s}\Big\rangle -$$

$$2\sum_{s\le n} \langle \overline{g}_{s-1} - \mathrm{Proj}_C(\overline{g}_{s-1}), \chi_s \frac{\overline{g}_{s-1} - \mathrm{Proj}_C(\overline{g}_{s-1})}{\tilde{\chi}_s}\Big\rangle +$$

$$2\sum_{s\le n} \Big\langle \chi_s \frac{\mathrm{Proj}_C(\overline{g}_{s-1}) - \overline{g}_{s-1}}{\tilde{\chi}_s}, \frac{g_s - \mathrm{Proj}_C(\overline{g}_{s-1})}{\tilde{\chi}_s}\Big\rangle.$$

Define $f_n = \overline{g}_n - \text{Proj}_C(\overline{g}_n)$, $j(n) = \min\{m;\ \tilde{\chi}_m \geq n\}$ (as in the proof of Theorem 1) and $\tilde{f}_n = \tilde{f}_{j(n)}$.

To show that $f_n \to 0$ almost surely, we prove first that $\sum_n \langle f_{n-1}, \frac{f_{n-1}}{\tilde{\chi}_n}\rangle < \infty$. By (3.11) and due to (d), it is sufficient to show that the following three series converge: $\sum_n \|\chi_n \frac{g_n - \text{Proj}_C(\overline{g}_{n-1})}{\tilde{\chi}_n}\|^2$, $\sum_n \|\chi_n \frac{\text{Proj}_C(\overline{g}_{n-1}) - \overline{g}_{n-1}}{\tilde{\chi}_n}\|^2$, and $\sum_n \left\langle \chi_n \frac{\text{Proj}_C(\overline{g}_{n-1}) - \overline{g}_{n-1}}{\tilde{\chi}_n}, \frac{g_n - \text{Proj}_C(\overline{g}_{n-1})}{\tilde{\chi}_n} \right\rangle$.

By assumption the sequence $\{g_n\}$ is bounded by an L_2 function and therefore, the sequence $\{\overline{g}_n\}$ is also L_2-bounded. Furthermore, by the assumption that C respects L_2-boundedness, $\{\text{Proj}_C(\overline{g}_n)\}$ is L_2-bounded. Thus, $\sum_n \|\chi_n \frac{g_n - \text{Proj}_C(\overline{g}_{n-1})}{\tilde{\chi}_n}\|^2 = \sum_n \int \frac{g_{j_n} - \text{Proj}_C(\overline{g}_{j_{n-1}})}{n^2} d\mu < \infty$. The second series converges for a similar reason. As for the third series,

$$\sum_n \left\langle \chi_n \frac{\text{Proj}_C(\overline{g}_{n-1}) - \overline{g}_{n-1}}{\tilde{\chi}_n}, \frac{g_n - \text{Proj}_C(\overline{g}_{n-1})}{\tilde{\chi}_n} \right\rangle = \sum_n \int \frac{\left(\text{Proj}_C(\overline{g}_{j_{n-1}}) - \overline{g}_{j_{n-1}}\right)\left(g_{j_n} - \text{Proj}_C(\overline{g}_{j_{n-1}})\right)}{n^2} d\mu,$$

which converges by the Cauchy-Schwartz inequality. We therefore obtain that $\sum_n \langle f_{n-1}, \frac{f_{n-1}}{\tilde{\chi}_n}\rangle < \infty$.

Since $\sum_n \langle f_{n-1}, \frac{f_{n-1}}{\tilde{\chi}_n}\rangle = \sum_n \langle \tilde{f}_{n-1}, \frac{\tilde{f}_{n-1}}{n}\rangle$, we obtain that $\sum_n \frac{\|\tilde{f}_{n-1}\|^2}{n} d\mu < \infty$ and therefore $\sum_n \frac{\|\tilde{f}_n\|^2}{n} d\mu < \infty$. This proves (c) of Proposition 1. In order to use Proposition 1 it remains to show that $\tilde{f}_n$ is an average of an a.s. bounded sequence.

Let $h_n = n(\overline{g}_{j_n} - \text{Proj}_C(\overline{g}_{j_n})) - (n-1)(\overline{g}_{j_{n-1}} - \text{Proj}_C(\overline{g}_{j_{n-1}}))$. Obviously, $\tilde{f}_n$ is the average of $h_1, ..., h_n$. To show that $\{h_n\}$ is almost surely bounded let $r_n = \overline{g}_{j_n}$. We obtain, due to assumption (b) of the theorem, that $|r_n - r_{n-1}| \leq \frac{1}{n}B$ for some function B. Since C respects L_2-boundedness, there exists a function B' such that $|\text{Proj}_C(r_n) - \text{Proj}_C(r_{n-1})| \leq \frac{1}{n}B'$. Thus, h_n is an a.s. bounded sequence.

Proposition 1 now ensures that $\tilde{f}_n \to 0$ a.s. and therefore $f_n \to 0$ a.s. and the proof is complete. ■

4 Games with Payoffs in Large Spaces

4.1 The case of $\chi \equiv 1$

Recall that $\overline{Y}_n(h^n)$ is the average payoff up to stage n when the history is h^n. In this section we suppose that χ is constantly 1. Thus, $\overline{Y}_n(h^n) = \frac{\sum_{h^s < h^n} Y_s(h^s)}{n}$. If at stage $n+1$ and after the history h^n, player 1's action is i and player 2's action is j, then the stage payoff is (the random variable) $Y_{n+1}(h^n, i, j)$ defined on Ω. If the stage mixed action of player 1 is p and that of player 2 is q, we denote $Y_{n+1}(h^n, p, q) = \sum_{(i,j)\in S_1\times S_2} p(i)q(j)Y(h^n, i, j)$.

Theorem 5 *Suppose that in a two-player game there is $0 < \delta$ such that the payoffs satisfy $\|Y_n(h^n)\|^2 = O(n^{\frac{1}{3}-\delta})$, where $\|\cdot\|$ is the L_2-morm. Furthermore, suppose that for any point $f \in L_2$ and after every history h^n there exists a mixed action of player 1, $p \in \Delta(S_1)$, such that for any mixed action of player 2, $q \in \Delta(S_2)$,*

$$\Big\langle f - \mathrm{Proj}_C(f), Y_{n+1}(h^n, p, q) - \mathrm{Proj}_C(f)\Big\rangle = O(n^{-\frac{2}{3}-\delta}).$$

Then, the set C is approachable by player 1.

Proof.

Let σ_1 be the following strategy. At stage 1 or if $\overline{Y}_n(h^n) \in C$ play an arbitrary mixed action. Otherwise play the mixed action $p \in \Delta(S_1)$, such that $\langle \overline{Y}_n(h_n) - \mathrm{Proj}_C(\overline{Y}_n(h^n)), Y_{n+1}(h^n, p, q) - \mathrm{Proj}_C(\overline{Y}_n(h^n))\rangle = O(n^{-\frac{2}{3}-\delta})$ for any $q \in \Delta(S_2)$.

Fix a strategy σ_2 of player 2 and denote by $\lambda = \lambda_{(\sigma_1,\sigma_2)}$ the probability induced by (σ_1, σ_2) over $\mathcal{H}$. Consider the probability space $(\mathcal{H} \times \Omega, \lambda \times \mu)$. Let $\mathcal{F}_n$ be the σ-field in $\mathcal{H} \times \Omega$ generated by histories of length n.

For every $(h, \omega) \in \mathcal{H} \times \Omega$ define $g_n(h, \omega) = Y_n(h^n)(\omega)$, where h^n is the n-prefix of h. By assumption g_n satisfies condition (a) of Theorem 3.

As for condition (b), by the construction of σ_1,

$$E\Big(\Big\langle \overline{Y}_n - \mathrm{Proj}_C(\overline{Y}_n), Y_{n+1}(h^n, i, j) - \mathrm{Proj}_C(\overline{Y}_n)\Big\rangle \Big| \mathcal{F}_n\Big) = O(n^{-\frac{2}{3}-\delta}).$$

Thus,

$$E\Big(E\Big(\Big\langle \overline{Y}_n - \mathrm{Proj}_C(\overline{Y}_n), Y_{n+1}(h^n, i, j) - \mathrm{Proj}_C(\overline{Y}_n)\Big\rangle\Big|\mathcal{F}_n\Big)\Big) = O(n^{-\frac{2}{3}-\delta}).$$

Therefore, $\langle \overline{g}_{n-1} - \mathrm{Proj}_C(\overline{g}_{n-1}), g_n - \mathrm{Proj}_C(\overline{g}_{n-1})\rangle = O(n^{-\frac{2}{3}-\delta})$ for every n and $\sum_{i=2}^{n}(i-1)\langle \overline{g}_{i-1} - \mathrm{Proj}_C(\overline{g}_{i-1}), g_i - \mathrm{Proj}_C(\overline{g}_{i-1})\rangle = \sum_{i=2}^{n} O(i^{\frac{1}{3}-\delta}) = O(n^{\frac{4}{3}-\delta})$. Thus, condition (b) is also satisfied.

Theorem 3 ensures that $\overline{Y}_n$ approaches C, as desired. ■

Remark. Note that if $C = \{0\}$, one can use the strategy σ_1 employed in the previous proof to show approachability to C by Theorem 2. The case of $C = \{0\}$ is sufficient for the generation of extended normal numbers (see Lehrer (2001c)).

5 Extending Approachability Further

In this section we deal with general χ. For any h^n define $\tilde{\chi}_{n+1}(h^n) = \sum_{h^k < h^n} \chi(h^k)$. Thus, $\tilde{\chi}_{n+1}(h^n)$ is the random variable which indicates how many times along the history h^n a point $\omega \in \Omega$ was active. Note that if χ is always 1, then $\tilde{\chi}_{n+1}(h^n) = n + 1$.

Recall that if the pair of actions (i, j) is played at stage $n + 1$ after the history h^n, then $\overline{Y}_{n+1}(h^n, i, j) = \frac{\tilde{\chi}_n(h^{n-1})\overline{Y}_n(h^n) + \chi_{n+1}(h^n)Y_{n+1}(h^n,i,j)}{\tilde{\chi}_{n+1}(h^n)}$.

Theorem 6 *Suppose that in a two-player game the payoffs are bounded by an L_2 function. Furthermore, suppose that there exists a sequence $\{\varepsilon_n\}$ satisfying $\sum_n \varepsilon_n < \infty$ such that for any point $f \in L_2$ and after every history h^n there exists a mixed action of player 1, $p \in \Delta(S_1)$, such that for any mixed action of player 2, $q \in \Delta(S_2)$,*

$$\Big\langle \chi(h^n)\frac{\overline{Y}_n(h^n) - \mathrm{Proj}_C(\overline{Y}_n(h^n))}{\tilde{\chi}_{n+1}(h^n)}, Y_{n+1}(h^n, p, q) - \mathrm{Proj}_C(\overline{Y}_n(h^n))\Big\rangle \le \varepsilon_n.$$

Then, the set C is approachable by player 1.

Proof. As in the proof of Theorem 5 define σ_1, the strategy of player 1, as follows. At stage 1 or if $\overline{Y}_n(h^n) \in C$, play an arbitrary mixed action. Otherwise, play the mixed action $p \in \Delta(S_1)$, such that

$$\left\langle \chi(h^n)\frac{\overline{Y}_n(h^n) - \mathrm{Proj}_C(\overline{Y}_n(h^n))}{\tilde{\chi}_{n+1}(h^n)}, Y_{n+1}(h^n, p, q) - \mathrm{Proj}_C(\overline{Y}_n(h^n))\right\rangle \leq \varepsilon_n$$

for any $q \in \Delta(S_2)$.

Fix a strategy σ_2 of player 2 and denote $\lambda = \lambda_{(\sigma_1,\sigma_2)}$. Consider the probability space $(\mathcal{H} \times \Omega, \lambda \times \mu)$. As in the previous proof, let $\mathcal{F}_n$ be the σ-field in $\mathcal{H} \times \Omega$ generated by histories of length n. For every $(h, \omega) \in \mathcal{H} \times \Omega$ define $g_n(h, \omega) = Y(h^n)(\omega)$. Since all the payoffs $Y(h^n)(\omega)$ are bounded by an L_2 function, the sequence g_n satisfies condition (b) of Theorem 4. By definition $\overline{g}_n = \overline{Y}_n$ satisfies (c) of Theorem 4.

We show that $\overline{g}_n - \mathrm{Proj}_C(\overline{g}_n)$ converges almost surely to zero. In order to use Theorem 4, we need to prove that condition (d) of this theorem is also satisfied. By the construction of σ_1,

$$E\left(\left\langle \chi(h^n)\frac{\overline{Y}_n(h^n) - \mathrm{Proj}_C(\overline{Y}_n(h^n))}{\tilde{\chi}_{n+1}(h^n)}, Y_{n+1}(h^n, i, j) - \mathrm{Proj}_C(\overline{Y}_n(h^n))\right\rangle \middle| \mathcal{F}_n\right) \leq \varepsilon_n,$$

and thus

$$E\left(E\left(\left\langle \chi(h^n)\frac{\overline{Y}_n(h^n) - \mathrm{Proj}_C(\overline{Y}_n(h^n))}{\tilde{\chi}_{n+1}(h^n)}, Y_{n+1}(h^n, i, j) - \mathrm{Proj}_C(\overline{Y}_n(h^n))\right\rangle \middle| \mathcal{F}_n\right)\right) \leq \varepsilon_n.$$

So, $\limsup_n \sum_{s \leq n} \left\langle \overline{g}_{s-1} - \mathrm{Proj}_C(\overline{g}_{s-1}), \chi_s \frac{g_s - \mathrm{Proj}_C(\overline{g}_{s-1})}{\tilde{\chi}_s}\right\rangle \leq \sum_n \varepsilon_n < \infty$, which guarantees that condition (d) is also satisfied.

Theorem 4 ensures that $\overline{g}_n - \mathrm{Proj}_C(\overline{g}_n)$ converges $\lambda \times \mu-$almost surely to zero. Hence, C is approachable by player 1 as desired. ■

6 Approachability as an Extension of the Law of Large Numbers

A simple version of the strong law of large number states that if $X_1, X_2, ...$ is a sequence of uncorrelated random variables with bounded variances, then the average $\frac{X_1+...+X_n-(E(X_1)+...+E(X_n))}{n}$ converges to 0. The approachability principles presented in Section 3 are extensions of the law of large numbers in that they deal with convergence to a closed set rather than to a point.

Definition 5 *Let C be a closed set and $Y_1, Y_2, ...$ be a sequence in L_2. Then, Y_{n+1} is C*-negatively correlated with $\overline{Y}_n = \frac{Y_1+...+Y_n}{\tilde{\chi}_n}$, *where $\tilde{\chi}_n = \chi_1+...+\chi_n$, if*

$$\langle \chi_{n+1} \frac{\overline{Y}_n - \mathrm{Proj}_C(\overline{Y}_n)}{\tilde{\chi}_{n+1}}, Y_{n+1} - \mathrm{Proj}_C(\overline{Y}_n) \rangle \leq 0.$$

Let $X_1, X_2, ...$ be a sequence in L_2 and $\chi_n \in \{0,1\}$ with the property that $\chi_n = 0$ implies $X_n = 0$ for every $n = 1, 2, ...$. Set $\tilde{\chi}_n = \chi_1 + ... + \chi_n$ and define $Y_n = X_n - E(X_n|\chi_n \neq 0)$ and $\overline{Y}_n = \frac{Y_1+...+Y_n}{\tilde{\chi}_n}$.

Note that if $X_1, X_2, ...$ is a sequence of uncorrelated random variables, then Y_{n+1} is $\{0\}$-negatively correlated with $\overline{Y}_n$, where $\chi_n = 1$ for every n. Corollary 2 implies that if $\chi_n = 1$ for every n, the variances of $Y_1, Y_2, ...$ are uniformly bounded and, finally, if for any n Y_{n+1} is $\{0\}$-negatively correlated with $\overline{Y}_n$, then $\overline{Y}_n$ converges almost surely to 0, that is, $\frac{X_1+...+X_n-(E(X_1|\chi_1\neq 0)+...+E(X_n|\chi_n\neq 0))}{\tilde{\chi}_n}$ converges almost surely to 0.

Theorems 3 and 4 provide necessary conditions for the convergence of the average $\overline{Y}_n$ to a closed set C. Theorem 3 implies that if $\chi_n = 1$ for every n, the variances of Y_n are uniformly bounded and Y_{n+1} is C-negatively correlated with $\overline{Y}_n$ for any n, then $\overline{Y}_n$ converges almost surely to C (i.e., $\overline{Y}_n - \mathrm{Proj}_C(\overline{Y}_n)$ converges almost surely to 0).

Theorem 4 implies that for general χ_n, if the sequence of Y_n is L_2-bounded, C respects L_2-boundedness, $\chi_n \to \infty$ almost surely and for any

n, Y_{n+1} is C-negatively correlated with $\overline{Y}_n$, then $\overline{Y}_n$ converges almost surely to C.

7 Final Remarks and Comments

7.1 Conjecture

I conjecture that the results stated in the first theorems can be sharpened. In Theorem 2, for instance, it seems sufficient to require that the sequence $\{g_n\}$ satisfies $\sum_n \frac{\|g_n\|^2}{n^2} < \infty$ in order to have convergence of $\{f_n\}$ to zero.

Also, the condition that the set C respects L_2-boundedness seems to be too strong. It is conjectured that if C is convex, for instance, then the conditions of Theorem 4 are sufficient to ensure approachability.

7.2 About the speed of convergence

In the case of general χ the proof of Theorem 1 implies that if $\langle \frac{f_{n-1}}{\tilde{\chi}_n}, g_n \rangle \leq 0$ for every n then $\sum_{n=1}^{\infty} \frac{\|\widetilde{f}_n\|^2}{n} \leq \sum_{n=1}^{\infty} \frac{c}{n^2}$ for some constant c. This fact indicates something about the rate $\|\widetilde{f}_n\|$ tends to zero. Is it true that $\|\widetilde{f}_n\| = O(n^{-\ frac12})$? For $\tilde{\chi}_n = n$ this is known (this is a direct consequence of (3.8)).

7.3 A generalization of Theorem 1

In Proposition 1 it is assumed that $\tilde{\chi}_n - \tilde{\chi}_{n-1}$ is either 0 or 1. It turns out that this proposition can be extended so as to allow $\tilde{\chi}_n - \tilde{\chi}_{n-1}$ to take any positive number, providing that these numbers are not too large compared to $\tilde{\chi}_n$. This is formally stated in what follows.

Theorem 1* *Suppose that*

(a) $\{\tilde{\chi}_n\}_0^{\infty}$ *is a sequence of non-decreasing random variables such that* $\tilde{\chi}_n \to \infty$ *μ-a.s. and* $\tilde{\chi}_0 = 0$*;*

(b) $\{g_n\}$ *is a sequence of* L_2*-bounded random variables s.t.* $\tilde{\chi}_n - \tilde{\chi}_{n-1} = 0$ *implies* $g_n = 0$*;*

(c) $f_n = \frac{\tilde{\chi}_{n-1} f_{n-1} + (\tilde{\chi}_n - \tilde{\chi}_{n-1}) g_n}{\tilde{\chi}_n}$*;*

(d) $\sum_n \langle \frac{(\tilde{\chi}_n - \tilde{\chi}_{n-1}) f_{n-1}}{\tilde{\chi}_n}, g_n \rangle < \infty$*; and*

(e) $\sum \| \frac{\tilde{\chi}_n - \tilde{\chi}_{n-1}}{\tilde{\chi}_n} \|^2 < \infty$*.*

Then, f_n *converges to zero with* μ*-probability 1.*

All the results that follow Theorem 1 can be extended in the same spirit and in particular the results in Sections 4 and 5. Suppose that the payoffs can be assigned weights greater than 1. The weights at stage $n+1$, after the history h^n, are given by the random variable $\tilde{\chi}(h^n) - \tilde{\chi}(h^{n-1})$. This random variable may attain any positive number, providing that $\sum_n E\Big(\| \frac{\tilde{\chi}(h^n) - \tilde{\chi}(h^{n-1})}{\tilde{\chi}(h^n)} \|^2\Big)$ is finite. Under these conditions Theorems 4 and 5 can be generalized in a natural way.

7.4 An analogous extension of the Ergodic Theorem

As demonstrated in Section 6, the geometric principles of approachability are extensions of the strong law of large numbers (that deal with uncorrelated random variables) to a convergence to a set, rather than to a point. What would be an appropriate extension of the Ergodic Theorem (that deals with stationary processes) and what then would be the related random variable-payoffs game?

8 References

Aumann, R.J. and M. Maschler with the collaboration of R. B. Streans (1995) *Repeated Games with Incomplete Information*, M.I.T Press.

Blackwell, D. (1956a) "Controlled Random Walks," in *Proceedings of the International Congress of Mathematicians*, Vol 3, 336-338, Amsterdam, North Holland, 1954.

Blackwell, D. (1956b) "An Analog of the MinMax Theorem For Vector Payoffs," *Pacific J. of Math.,* **6**, 1-8.

David, A. P. (1982) "The Well Calibrated Bayesian," *Journal of the American Statistical Association,* **77**, 379, 605-613.

Foster, D., and R. Vohra (1997) "Calibrated Learning and Correlated Equilibrium," *Games and Economic Behavior*, **21**, 40-55.

Foster, D., and R. Vohra (1999) "Regret in the On-Line Decision Problem," *Games and Economic Behavior*, **29**, 7-35.

Hannan, J. (1957) "Approximation to Bayes Risk in Repeated Plays," in *Contribution to the Theory of Games*, **3**, 97-139. Princeton, NJ: Princeton University Press.

Hart, S., and A. Mas-Colell (2000) "A Simple Adaptive Procedure Leading to Correlated Equilibrium," *Econometrica*, **68**, 1127-1150.

Hart, S., and A. Mas-Colell (2001) "A General Class of Adaptive Strategies," *Journal of Economic Theory*, **98**, 26-54.

Kuipers, L. and H. Niederreiter (1974) *Uniform Distribution of Sequences*, John Wiley & Sons, New York.

Lehrer, E. (2001a) "Any Inspection is Manipulable," *Econometrica*, **69**, 5, 1333-1347. Can also be found in www.math.tau.ac.il/~lehrer/Papers

Lehrer, E. (1997/2001b) "A Wide Range no-Regret Theorem ," to appear in *Games and Economic Behavior*; www.math.tau.ac.il/~lehrer/Papers

Lehrer, E. (1997/2001c) "The Game of Normal Numbers," mimeo; www.math.tau.ac,il/~lehrer/Papers

Mertens, J.-F, S. Sorin and S. Zamir (1994) "Repeated Games," CORE discussion paper #9420.

Rustichini, A. (1999) "Minimizing Regret: The General Case," *Games and Economic Behavior*, **29**, 224-243.

Sandroni, A., R. Smorodinsky, and R. Vohra (2000) "Calibration with Multiple Checking Rules," mimeo.

Spinat, X. (2000) "A Necessary and Sufficient Condition for Approachability," *Mathematics of Operations Research*, forthcoming.

Vieille, N (1992) "Weak Approachability," *Mathematics of Operations Research*, **17** (4), 781-792.