14 Repeated games with vector payoffs

Chapter summary

This chapter is devoted to a theory of repeated games with vector payoffs, known as the theory of *approachability*, developed by Blackwell in 1956. Blackwell considered two-player repeated games in which the outcome is an m-dimensional vector of attributes, and the goal of each player is to control the average vector of attributes. The goal can be either to *approach* a given target set $S \subseteq \mathbb{R}^m$, that is, to ensure that the distance between the vector of average attributes and the target set S converges to 0, or to *exclude* the target set S, that is, to ensure that the distance between the vector of average attributes and S remains bounded away from 0. If a player can approach the target set we say that the set is *approachable* by the player, whereas if the player can exclude the target set we say that it is *excludable* by that player. Clearly, a set cannot be both approachable by one player and excludable by the other player.

We provide a geometric condition that ensures that a set is approachable by a player, and show that any convex set is either approachable by one player or excludable by the other player.

Two applications of the theory of approachability are provided: it is used, respectively, to construct an optimal strategy for the uninformed player in two-player zero-sum repeated games with incomplete information on one side, and to construct a no-regret strategy in sequential decision problems with experts.

In Chapter 13 we studied repeated games in which the payoff to each player in every stage was a real number representing the player's utility. In this chapter we will look at two-player repeated games in which the outcome in every stage is not a pair of payoffs, but a vector in the m-dimensional Euclidean space $\mathbb{R}^m$. These games correspond to situations in which the outcome of an interaction between the players is comprised of several incommensurable factors. For example, an employment contract between an employee and an employer may specify the number of hours the employee is to commit to the job; the salary the employee will receive; and the number of days of annual leave granted to the employee. As we saw in Chapter 2 on utility theory, under certain assumptions it is possible to associate each outcome with a real number representing the utility of the outcome, thereby translating the situation into a game with payoffs in real numbers. But we may not know the players' utility functions. In addition, we may at times be interested in controlling each variable separately, as is done for example in physics problems, where pressure and temperature may be controlled separately. The model of repeated games with vector payoffs was first presented by Blackwell [1956]. The first part of this chapter is based on that paper.

When the outcome of an interaction to each player is a payoff, each player tries to maximize the average of the payoffs he receives. When the outcome is a vector in $\mathbb{R}^m$, maximizing one coordinate may come at the expense of another coordinate. We therefore speak of target sets in the space of vector payoffs: each player tries either to cause the average of his payoffs to approach a target set (i.e., a certain subset of $\mathbb{R}^m$) or to exclude a target set.

In Chapters 9 and 10 we studied Bayesian games; these are games with incomplete information whose payoffs depend on the state of nature, which can have a finite number of values. In Section 14.7 (page 590) we will study two-player zero-sum repeated games with incomplete information regarding the state of nature using the model of repeated games with vector payoffs: every pair of actions in such a game is associated with a vector of payoffs composed of the payoff for each possible state of nature. In this way, we can monitor the average payoff for every possible state of nature, even if the state of nature is not known by all the players. An example of such an application appears in Section 14.7 (page 590). In Section 14.8 (page 600), we will present an additional application of the model of repeated games with vector payoffs to the study of dynamic decision problems with experts.

14.1 Notation

In this chapter we will work in $\mathbb{R}^m$, the m-dimensional Euclidean space. We will sometimes term $x \in \mathbb{R}^m$ a "vector," and sometimes a "point." The zero vector in $\mathbb{R}^m$ is denoted by $\vec{0}$.

Recall that for a finite set A, we denote by $\Delta(A)$ the set of probability distributions over A. The inner product in $\mathbb{R}^m$ is denoted as follows. For every pair of vectors $x, y \in \mathbb{R}^m$,

$$\langle x, y\rangle := \sum_{l=1}^{m} x_l y_l. \tag{14.1}$$

The inner product is symmetric, $\langle x, y\rangle = \langle y, x\rangle$, and *bilinear*; i.e., it is a linear function in each of its variables. That is, for every $\alpha, \beta \in \mathbb{R}$ and every $x, x_1, x_2, y, y_1, y_2 \in \mathbb{R}^m$,

$$\langle \alpha x_1 + \beta x_2, y\rangle = \alpha\langle x_1, y\rangle + \beta\langle x_2, y\rangle, \tag{14.2}$$

and

$$\langle x, \alpha y_1 + \beta y_2\rangle = \alpha\langle x, y_1\rangle + \beta\langle x, y_2\rangle. \tag{14.3}$$

The norm of a vector $x \in \mathbb{R}^m$, denoted by $\|x\|$, is the Euclidean norm, given by

$$\|x\| := \langle x, x\rangle^{1/2} = \sqrt{\sum_{l=1}^{m} (x_l)^2}, \tag{14.4}$$

and the distance function between vectors is

$$d(x, y) := \|x - y\| = \langle x - y, x - y\rangle^{1/2} = \sqrt{\sum_{l=1}^{m}(x_l - y_l)^2}. \qquad (14.5)$$

If $C \subseteq \mathbb{R}^m$ is a set, and $x \in \mathbb{R}^m$ is a vector, the distance between x and C is given by

$$d(x, C) := \inf_{y \in C} d(x, y). \qquad (14.6)$$

It follows that the distance between a point x and a set C equals the distance between x and the closure of C, and $d(x, C) = 0$ for every x in the closure of C. The triangle inequality states that

$$d(x, y) + d(y, z) \geq d(x, z), \quad \forall x, y, z \in \mathbb{R}^m. \qquad (14.7)$$

Equivalently,

$$\|x\| + \|y\| \geq \|x + y\|. \qquad (14.8)$$

The Cauchy–Schwartz inequality states that

$$\|x\|^2 \|y\|^2 \geq \langle x, y\rangle^2. \qquad (14.9)$$

The following inequalities also hold (Equation (14.11) follows from the Cauchy–Schwartz inequality):[1]

$$d(x + y, x + z) = d(y, z), \qquad \forall x, y, z \in \mathbb{R}^m, \qquad (14.10)$$
$$d(x + y, z + w) \leq d(x, z) + d(y, w), \qquad \forall x, y, z, w \in \mathbb{R}^m, \qquad (14.11)$$
$$d(\alpha x, \alpha y) = \alpha d(x, y), \qquad \forall x, y \in \mathbb{R}^m, \forall \alpha > 0, \qquad (14.12)$$
$$d(x, y) \leq 2M\sqrt{m}, \qquad \forall M > 0, \forall x, y \in [-M, M]^m. \qquad (14.13)$$

If $C \subseteq \mathbb{R}^m$ is a set, and $x, y \in \mathbb{R}^m$ are vectors, then (Exercise 14.1)

$$d(x, C) \leq d(x, y) + d(y, C). \qquad (14.14)$$

All the vectors are considered to be row vectors. If x is a row vector, then $x^\top$ is the corresponding column vector.

Since we are studying two-player games, for every player $k \in \{1, 2\}$, we will denote by $-k$ the player who is not player k. In particular, the notation σ_{-k} denotes a strategy of the player who is not k.

---

[1] For every set $A \subseteq \mathbb{R}$, and natural number m, the set $A^m \subseteq \mathbb{R}^m$ is defined as follows:

$$A^m = \underbrace{A \times A \times \cdots \times A}_{m \text{ times}} = \{(x_1, x_2, \ldots, x_m) \in \mathbb{R}^m : x_i \in A, \quad i = 1, 2, \ldots, m\}.$$

14.2 The model

Definition 14.1 *A* repeated (two-player) game with (m-dimensional) vector payoffs *is given by two action sets* $\mathcal{I} = \{1, 2, \ldots, I\}$ *and* $\mathcal{J} = \{1, 2, \ldots, J\}$ *of Players* 1 *and* 2*, respectively,*[2] *and a payoff function* $u : \mathcal{I} \times \mathcal{J} \to \mathbb{R}^m$.

As previously stated, the vectors in $\mathbb{R}^m$ are not necessarily payoffs; they are various attributes of the outcome of the game. Despite this, we use the term "payoff function" for u, both for convenience and because of the analogy to games with scalar payoffs (the case $m = 1$). It will sometimes be convenient to present the payoff function u as a matrix of order $I \times J$, whose elements are vectors in $\mathbb{R}^m$.

The game proceeds in stages as follows. In stage t, ($t = 1, 2, \ldots$), each one of the players chooses an action: Player 1 chooses action $i^t \in \mathcal{I}$, and Player 2 chooses action $j^t \in \mathcal{J}$. As in the model of repeated games, we will assume that every player knows what the other player chose in previous stages. A behavior strategy of Player 1 is a function associating a mixed action with each history of actions

$$\sigma_1 : \bigcup_{t=1}^{\infty} (\mathcal{I} \times \mathcal{J})^{t-1} \to \Delta(\mathcal{I}). \tag{14.15}$$

Similarly, a behavior strategy of Player 2 is a function

$$\sigma_2 : \bigcup_{t=1}^{\infty} (\mathcal{I} \times \mathcal{J})^{t-1} \to \Delta(J). \tag{14.16}$$

Kuhn's Theorem for infinite games (Theorem 6.26 on page 242) states that every mixed strategy has an equivalent behavior strategy and vice versa. It therefore suffices to consider only behavior strategies here, because they are more natural than mixed strategies. The word *strategy* in this chapter will be short-hand for "behavior strategy."

By Theorem 6.23 (page 242), every pair of strategies (σ_1, σ_2) induces a probability measure $\mathbf{P}_{\sigma_1,\sigma_2}$ over the set of infinite plays, i.e., over $(\mathcal{I} \times \mathcal{J})^{\mathbb{N}}$. The expectation operator corresponding to this probability distribution is denoted by $\mathbf{E}_{\sigma_1,\sigma_2}$.

Denote the payoff in stage t by $g^t = u(i^t, j^t) \in \mathbb{R}^m$, and the average payoff up to stage T by[3]

$$\overline{g}^T = \frac{1}{T} \sum_{t=1}^{T} g^t = \frac{1}{T} \sum_{t=1}^{T} u(i^t, j^t) \in \mathbb{R}^m. \tag{14.17}$$

We next define the concept of an approachable set, the central concept of this chapter.

---

[2] For convenience, we use in this chapter the notation $\mathcal{I}$ and $\mathcal{J}$ for the action sets of the players, instead of A_1 and A_2.

[3] While in one-stage games the payoff is defined to be the expected payoff according to the mixed actions of the players, in repeated games the payoff in each stage t is the actual payoff $u(i^t, j^t)$ of that stage (and not the expected payoff according to the mixed actions at that stage). In this chapter we will be interested in the average payoff $\overline{g}^T$, as opposed to its expectation.

Definition 14.2 *A nonempty set* $C \subseteq \mathbb{R}^m$ *is called* approachable *by player* k *if there exists a strategy* σ_k *of player* k *such that for every* $\varepsilon > 0$ *there exists* $T \in \mathbb{N}$ *such that for every strategy* σ_{-k} *of the other player*

$$\mathbf{P}_{\sigma_k,\sigma_{-k}}(d(\overline{g}^t, C) < \varepsilon, \quad \forall t \geq T) > 1 - \varepsilon. \tag{14.18}$$

In this case we say that σ_k approaches C *for player* k.

A set is approachable by a player if that player can guarantee that for any strategy used by the other player, the average payoff approaches the set with probability 1 uniformly. In particular, this implies that

$$\mathbf{P}_{\sigma_k,\sigma_{-k}}(\lim_{t\to\infty} d(\overline{g}^t, C) = 0) = 1. \tag{14.19}$$

The convergence of the average payoff to C is uniform; i.e., the rate at which the average payoff approaches this set (meaning the ratio between ε and t in Equation (14.18)) is independent of the strategy used by the rival player.

The dual to Definition 14.2 relates to the situation in which player k can guarantee that the distance between the average payoff and the target set is positive and bounded away from 0.

Definition 14.3 *A nonempty set* $C \subseteq \mathbb{R}^m$ *is called* excludable *by player* k *if there exists* $\delta > 0$ *such that the set* $\{x \in \mathbb{R}^m : d(x, C) \geq \delta\}$ *is approachable by player* k. *If the strategy* σ_k *of player* k *approaches the set* $\{x \in \mathbb{R}^m : d(x, C) \geq \delta\}$ *for some* $\delta > 0$, *we say that* σ_k excludes *the set* C *for player* k.

14.3 Examples

When $m = 1$, the outcome at each stage is a real number. If we interpret this number as the payoff to Player 1, and the negative of this number as the payoff to Player 2, then this model is equivalent to the model of repeated two-player zero-sum games. If v is the value of the one-stage game, then $[v, \infty)$ is an approachable set for Player 1, and $(-\infty, v]$ is an approachable set for Player 2. The players' approaching strategies are stationary strategies, in which each player plays an optimal strategy of the one-stage game at each stage (independently of the history of play). It follows that for every $\delta > 0$, the set $(-\infty, v - \delta]$ is an excludable set for Player 1, and the set $[v + \delta, \infty)$ is an excludable set for Player 2. This example shows that one may regard the model of repeated games with vector payoffs as a generalization of the model of two-player zero-sum games. Blackwell [1956], in fact, presented his model in such a way.

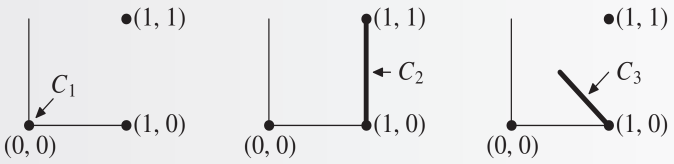
Example 14.4 Consider a game where $m = 2$, each player has two possible actions, and the payoff function u is given by the matrix in Figure 14.1.

		Player 2	
		L	R
Player 1	T	(0, 0)	(0, 0)
	B	(1, 1)	(1, 0)

Figure 14.1 The game in Example 14.4

The set $C_1 = \{(0, 0)\}$, containing only the vector (0, 0) (see Figure 14.2), is approachable by Player 1: if Player 1 plays T in every stage, he guarantees that the average payoff is (0, 0).

The set $C_2 = \{(1, x) \colon 0 \leq x \leq 1\}$ (see Figure 14.2) is also approachable by Player 1: if Player 1 plays B in every stage, he guarantees that the average payoff is in C_2.



The set C_1 The set C_2 The set C_3

Figure 14.2 Three sets approachable by Player 1 in Example 14.4

It is also interesting to note that the set $C_3 = \left\{(x, 1 - x) \colon \frac{1}{2} \leq x \leq 1\right\}$ (see Figure 14.2) is also approachable by Player 1. The following strategy of Player 1 guarantees that the average payoff approaches this set:

- If $\overline{g}^{t-1}$, the average payoff up to stage $t - 1$, is located above the diagonal $x_1 + x_2 = 1$, i.e., if $\overline{g}_1^{t-1} + \overline{g}_2^{t-1} \geq 1$, then play T in stage t.
- If $\overline{g}^{t-1}$, the average payoff up to stage $t - 1$, is located below the diagonal $x_1 + x_2 = 1$, i.e., if $\overline{g}_1^{t-1} + \overline{g}_2^{t-1} < 1$, then play B in stage t.

In Exercise 14.8 we present a guided proof of the fact that the set C_3 is indeed approachable by Player 1. ◀

14.4 Connections between approachable and excludable sets

The following claims, whose proofs are left to the reader, state several simple properties that follow from the definitions (Exercise 14.4).

Theorem 14.5 *The following two claims hold:*

1. *If strategy σ_k approaches a set C for player k, then it approaches the closure of C for that player.*

2. *If strategy σ_k excludes a set C for player k, then it excludes the closure of C for that player.*

Let $M \geq \frac{1}{2}$ be a bound on the norm of the payoffs in the game

$$\|u(i, j)\| \leq M, \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J}. \tag{14.20}$$

In particular, $\|u(i^t, j^t)\| \leq M$, in every stage t. The triangle inequality implies that

$$\|\overline{g}^T\| \leq M, \quad \forall T \in \mathbb{N}. \tag{14.21}$$

In words, the average payoff is located in the ball with radius M around the origin. Therefore, if the average payoff approaches a particular set, it must approach the intersection of that set and the ball of radius M around the origin. Similarly, if a player can guarantee that the distance between the average payoff and a particular set is positive and bounded away from 0, then he can guarantee that the distance between the average payoff and the intersection of that set and the ball of radius M around the origin is positive and bounded away from 0. This insight is expressed in the next theorem, whose proof is left to the reader (Exercise 14.5).

Theorem 14.6 *The following two claims hold:*

1. *A closed set C is approachable by a player if and only if the set $\{x \in C : \|x\| \leq M\}$ is approachable by the player.*
2. *A closed set C is excludable by a player if and only if the set $\{x \in C : \|x\| \leq M\}$ is excludable by the player.*

The following theorem relates to sets containing approachable sets, and to subsets of excludable sets (Exercise 14.6).

Theorem 14.7 *The following two claims hold:*

1. *If strategy σ_k approaches a set C for player k, then it approaches every superset of C for that player.*
2. *If strategy σ_k excludes a set C for player k, then it excludes every subset of C for that player.*

We close this section with the following theorem (Exercise 14.7).

Theorem 14.8 *A set C cannot be both approachable by one player and excludable by the other player.*

Theorem 14.8 expresses the opposing interests of the players in this model, as in the model of two-player zero-sum games. In the next section we will present a geometric condition for the approachability of a set, which we then use to prove that every closed and convex set is either approachable by one player, or excludable by the other player.

14.5 A geometric condition for the approachability of a set

If in stage t Player 1 plays the mixed action p and Player 2 plays the mixed action q, the expected payoff in that stage is[4]

$$U(p,q) := \sum_{i,j} p_i u(i,j) q_j, \tag{14.22}$$

which is a vector in $\mathbb{R}^m$. For every mixed action $p \in \Delta(\mathcal{I})$ of Player 1, define the set

$$R_1(p) := \{U(p,q) \colon q \in \Delta(\mathcal{J})\} = \left\{ \sum_{i,j} p_i u(i,j) q_j \colon q \in \Delta(\mathcal{J}) \right\} \subseteq \mathbb{R}^m. \tag{14.23}$$

Thus, if Player 1 plays the mixed action p, the expected payoff in the current stage is in the set $R_1(p)$. As we will show (Theorem 14.19, page 585), for every $p \in \Delta(\mathcal{I})$, the strategy of Player 1 in which he plays the mixed action p in every stage approaches the set $R_1(p)$. The reason for this is that when Player 1 implements the mixed action p in every stage, the expected payoff in each stage is located in $R_1(p)$, independently of the action implemented by Player 2. Since the set $R_1(p)$ is convex, it follows that for every $T \in \mathbb{N}$ the expectation of the average payoff up to stage T is also in $R_1(p)$. As we will later show this further implies, by way of a variation of the strong law of large numbers, that the average payoff $\bar{g}^T$ approaches $R_1(p)$ as T increases to infinity.

Similarly, for every mixed action $q \in \Delta(\mathcal{J})$ of Player 2, defines

$$R_2(q) := \{U(p,q) \colon p \in \Delta(\mathcal{I})\} = \left\{ \sum_{i,j} p_i u(i,j) q_j \colon p \in \Delta(\mathcal{I}) \right\} \subseteq \mathbb{R}^m. \tag{14.24}$$

Just as for $R_1(p)$, for every $q \in \Delta(\mathcal{J})$, the strategy of Player 2 in which he plays the mixed action q in every stage approaches the set $R_2(q)$ (Theorem 14.19, page 585).

Example 14.9 Consider the game with two-dimensional payoffs in Figure 14.3.

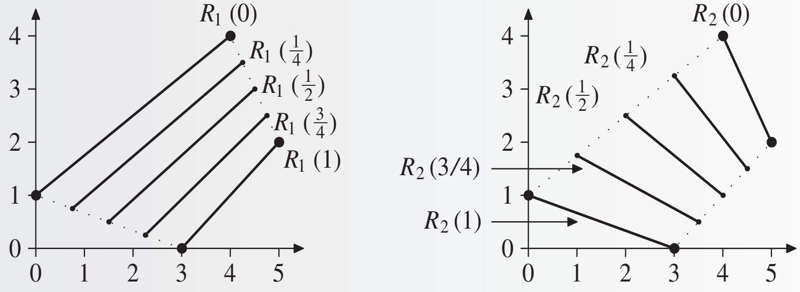
		Player 2	
		L	R
Player 1	T	(3, 0)	(5, 2)
	B	(0, 1)	(4, 4)

Figure 14.3 The game in Example 14.9

---

[4] Here, and in the rest of this chapter, a sum $\sum_{i,j}$ will be understood to mean the double sum $\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}}$.

Figure 14.4 depicts the sets $R_1(p)$ and $R_2(q)$ for several values of p and q. For simplicity, when Player 1 has two actions, T and B, we will identify every number p in the interval $[0, 1]$ with the mixed action $[p(T), (1-p)(B)]$. When Player 2 has two actions, L and R, we will identify every number q in the interval $[0, 1]$ with the mixed action $[q(L), (1-q)(R)]$.



The sets $R_1\,(p)$ The sets $R_1\,(p)$

Figure 14.4 The sets $R_1(p)$ and $R_2(p)$ in Example 14.9

◀

Definition 14.10 *A* hyperplane $H(\alpha, \beta)$ *in* $\mathbb{R}^m$ *is defined by*

$$H(\alpha, \beta) := \left\{x \in \mathbb{R}^m : \langle \alpha, x\rangle = \beta\right\}, \tag{14.25}$$

where $\alpha \in \mathbb{R}^m$ *and* $\beta \in \mathbb{R}$.

Denote

$$H^+(\alpha, \beta) = \{x \in \mathbb{R}^m : \langle x, \alpha\rangle \geq \beta\} \tag{14.26}$$

and

$$H^-(\alpha, \beta) = \{x \in \mathbb{R}^m : \langle x, \alpha\rangle \leq \beta\}. \tag{14.27}$$

$H^+(\alpha, \beta)$ and $H^-(\alpha, \beta)$ are the *half-spaces* defined by the hyperplane $H(\alpha, \beta)$. Note that $H^+(\alpha, \beta) \cap H^-(\alpha, \beta) = H(\alpha, \beta)$. Figure 14.5 depicts the hyperplane $H((2, 1), 2)$ in $\mathbb{R}^2$, and the two corresponding half-spaces.

By definition (see Corollary 14.23),

$$H^+(\alpha, \beta) = H^-(-\alpha, -\beta). \tag{14.28}$$

For every $x, y \in \mathbb{R}^m$, the hyperplane $H(x-y, \langle x-y, y\rangle)$ is the hyperplane passing through the point y, and perpendicular to the line passing through x and y (Exercise 23.35 on page 954). For example, in the case $m = 2$ described in Figure 14.6, the slope of the line passing through x and y is $\frac{y_2-x_2}{y_1-x_1}$. We now show that the slope of the hyperplane $H(x-y, \langle x-y, y\rangle)$, which in this case is a line, is $-\frac{y_1-x_1}{y_2-x_2}$, and therefore this line is perpendicular to the line passing through x and y. Choose a point $z = (z_1, z_2) \neq y$ on the hyperplane $H(x-y, \langle x-y, y\rangle)$. Then z satisfies

$$z_1(x_1-y_1) + z_2(x_2-y_2) = \langle x-y, y\rangle = (x_1-y_1)y_1 + (x_2-y_2)y_2. \tag{14.29}$$

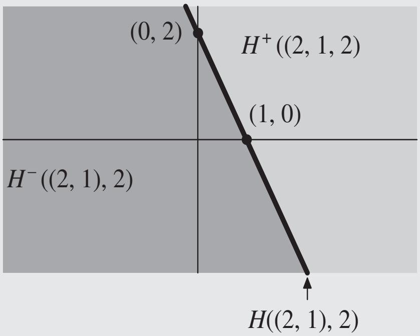


Figure 14.5 The hyperplane $H((2, 1), 2)$ in $\mathbb{R}^2$, which is the line $2x_1 + x_2 = 2$

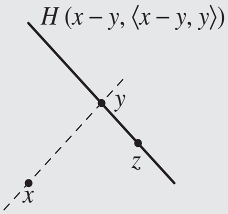


Figure 14.6 The hyperplane $H(y - x, \langle y - x, y\rangle)$

This further implies that the slope of the line connecting z and y, which is the hyperplane H, is

$$\frac{z_2 - y_2}{z_1 - y_1} = -\frac{x_1 - y_1}{x_2 - y_2}, \tag{14.30}$$

which is what we needed to show.

Definition 14.11 *Let $C \subseteq \mathbb{R}^m$ be a set, and let $x \notin C$ be a point in $\mathbb{R}^m$. A hyperplane $H(\alpha, \beta)$ is said to* separate *x from C if:*

1. *$x \in H^+(\alpha, \beta) \setminus H(\alpha, \beta)$ and $C \subseteq H^-(\alpha, \beta)$, or*
2. *$x \in H^-(\alpha, \beta) \setminus H(\alpha, \beta)$ and $C \subseteq H^+(\alpha, \beta)$.*

In words, a hyperplane $H(\alpha, \beta)$ separates x from C if (i) $\langle x, \alpha\rangle > \beta$ and $\langle y, \alpha\rangle \leq \beta$ for all $y \in C$, or (ii) $\langle x, \alpha\rangle < \beta$ and $\langle y, \alpha\rangle \geq \beta$ for all $y \in C$.

As in Chapter 13, denote by F the convex hull of all possible one-stage payoffs:

$$F = \text{conv}\{u(i, j), (i, j) \in \mathcal{I} \times \mathcal{J}\}. \tag{14.31}$$

Note that the average payoff $\overline{g}^t$, as a weighted average of vectors in the convex set $\{u(i, j), (i, j) \in \mathcal{I} \times \mathcal{J}\}$, is necessarily in the set F.

As previously noted, it will follow that the set $R_1(p)$ will be proved to be approachable by Player 1, for every $p \in \Delta(\mathcal{I})$. By Theorem 14.7, any half-space containing at least one of the sets $(R_1(p))_{p\in\Delta(\mathcal{I})}$ is also approachable by Player 1. This observation leads to the concept of a "B-set." A set C is a B-set for Player 1 if each half-space in a certain collection of half-spaces contains a set $R_1(p)$ for some $p \in \Delta(\mathcal{I})$.

Definition 14.12 *A closed set $C \subseteq \mathbb{R}^m$ is a B-set for Player 1 if for every point $x \in F \setminus C$ there exist a point $y = y(x, C) \in C$ and a mixed action $p = p(x, C) \in \Delta(\mathcal{I})$ of Player 1 satisfying:*

1. *y is a point in C that is closest to x:*

$$d(x, y) = d(x, C). \tag{14.32}$$

2. *The hyperplane $H(y - x, \langle y - x, y\rangle)$ separates x from $R_1(p)$:*

$$R_1(p) \subseteq H^+(y - x, \langle y - x, y\rangle), \tag{14.33}$$

$$x \in H^-(y - x, \langle y - x, y\rangle) \setminus H(y - x, \langle y - x, y\rangle). \tag{14.34}$$

Remark 14.13 *The hyperplane $H(y - x, \langle y - x, y\rangle)$ satisfies the following three properties (Exercise 23.35 on page 954):*

1. *$y \in H(y - x, \langle y - x, y\rangle)$.*
2. *This hyperplane is perpendicular to $y - x$, that is, $\langle y - x, z - y\rangle = 0$ for all $z \in H(y - x, \langle y - x, y\rangle)$.*
3. *y is the point in $H(y - x, \langle y - x, y\rangle)$ that is closest to x, that is, $\langle z - x, z - x\rangle > \langle y - x, y - x\rangle$ for all $z \in H(y - x, \langle y - x, y\rangle)$, $z \neq y$.*

Similarly, for a given hyperplane H and a point $x \notin H$, if $y \in H$ is the point in H that is closest to x, then $H = H(y - x, \langle y - x, y\rangle)$ (Exercise 23.36 on page 954). ◆

Note that the condition in Definition 14.12 requires that for every x there exist a point y and a mixed action p of Player 1 satisfying (a) and (b); for a given mixed action p, it is not the case that every point y satisfying (a) also satisfies (b). In Figure 4.7, there are two points y, y' in C that are the closest points to x. The hyperplane $H(y - x, \langle y - x, y\rangle)$, containing y, separates x from $R_1(p)$. In contrast, the hyperplane $H(y' - x, \langle y' - x, y'\rangle)$ does not separate x from $R_1(p)$.

The definition of a B-set for Player 2 is analogous to Definition 14.12: a set C is a B-set for Player 2 if for each point $x \in F \setminus C$ there exists a mixed action $q \in \Delta(\mathcal{J})$ of Player 2 such that the hyperplane $H(y - x, \langle y - x, y\rangle)$ separates x from $R_2(q)$, where $y \in C$ is a point in C that is closest to x. The following theorem presents a geometric condition that guarantees the approachability of a set by a particular player.

Theorem 14.14 (Blackwell [1956]) *If a set C is a B-set for player k, then it is approachable by player k.*

The converse may not hold: there are sets approachable by a player k that are not B-sets for player k (Exercise 14.15).

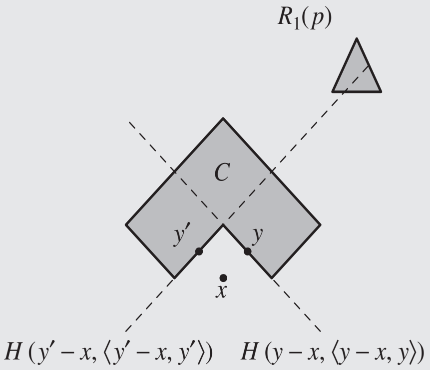


Figure 14.7 The hyperplane $H(y-x, \langle y-x, y\rangle)$ separates x from $R_1(p)$

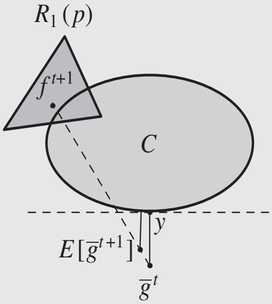


Figure 14.8 The idea behind the proof of Theorem 14.14

The intuition behind the proof (for Player 1) is depicted in Figure 14.8. Consider the strategy of Player 1 under which he plays the mixed action $p(\overline{g}^t, C)$ in every stage t. The hyperplane identified by the definition of a B-set is the hyperplane tangent to C at the point $y = y(\overline{g}^t, C)$ in C that is closest to x. Suppose that $\overline{g}^t$, the average payoff up to stage t is outside C, and let $p = p(\overline{g}^t, C)$ be the mixed action of Player 1, respectively, satisfying conditions (1) and (2) in Definition 14.12, for $x = \overline{g}^t$. If Player 1 plays the mixed action p, the expected payoff in stage $t+1$, denoted in the figure by f^{t+1}, is in $R_1(p)$, and therefore the expected value of $\overline{g}^{t+1}$ is located on the line connecting $\overline{g}^t$ with f^{t+1}. We will show that the expected distance $d(\overline{g}^{t+1}, C)$ is smaller than $d(\overline{g}^t, C)$; i.e., the expected distance between $\overline{g}^{t+1}$ and C is smaller than the distance between $\overline{g}^t$ and C. Finally, we will show that if the expected distance to C goes to 0, the distance itself also goes to 0, with probability 1.

We now turn to the formal proof of the theorem.

Proof: We will prove the theorem for Player 1. The proof for Player 2 is similar. From Theorem 14.6 (page 575) we may assume without loss of generality that for every $y \in C$,

$$\|y\| \leq M, \tag{14.35}$$

and, in particular, the absolute value of every coordinate of y is less than or equal to M. We will first define a strategy σ_1^* for Player 1, and then prove that it guarantees that the average payoff approaches the set C. In the first stage, the strategy σ_1^* chooses any action. For each $t \geq 1$ the strategy σ_1^* instructs Player 1 to play as follows in stage $t+1$:

- If $\overline{g}^t \in C$, the definition of σ_1^* is immaterial (play any action).
- If $\overline{g}^t \notin C$, the strategy σ_1^* instructs the player to choose the mixed action $p(\overline{g}^t, C)$ (as defined in Definition 14.12).

Denote by $d^t = d(\overline{g}^t, C)$ the distance between the average payoff up to stage t and the set C. We wish to show that for every strategy σ_2 of Player 2, the distance d^t converges to zero, with probability 1, and that the rate of convergence can be bounded, independently of the strategy of Player 2.

Lemma 14.15 *For every strategy σ_2 of Player 2, and for every $t \in \mathbb{N}$,*

$$\mathbf{E}_{\sigma_1^*,\sigma_2}[(d^t)^2] \leq \frac{4M^2}{t}. \tag{14.36}$$

Proof: We will prove the claim by induction on t. Since the payoffs are bounded by M, one has $d^t = d(\overline{g}^t, C) \leq 2M$ for all $t \in \mathbb{N}$: the distance between the average payoff and the set C is not greater than twice the maximal payoff. Since $M \geq \frac{1}{2}$, Equation (14.36) holds, for $t = 1$.

Assume by induction that Equation (14.36) holds for t; we will prove that it holds for $t+1$. The average payoff up to stage $t+1$ is a weighted average (i) of the average payoff up to stage t, and (ii) of the payoff in stage $t+1$:

$$\overline{g}^{t+1} = \frac{1}{t+1}\sum_{l=1}^{t+1} g^l = \frac{t}{t+1} \times \frac{1}{t}\sum_{l=1}^{t} g^l + \frac{1}{t+1}g^{t+1} = \frac{t}{t+1}\overline{g}^t + \frac{1}{t+1}g^{t+1}. \tag{14.37}$$

We wish to show that the expected value of d^{t+1}, the distance between $\overline{g}^{t+1}$ and C, shrinks. If $\overline{g}^t \in C$, then $y^t = \overline{g}^t$. If $\overline{g}^t \notin C$, denote by p^t the mixed action that Player 1 plays in stage $t+1$. Since $y^t \in C$, one has $d(\overline{g}^{t+1}, C) \leq d(\overline{g}^{t+1}, y^t)$, leading to

$$(d^{t+1})^2 = (d(\overline{g}^{t+1}, C))^2 \leq (d(\overline{g}^{t+1}, y^t))^2 = \|y^t - \overline{g}^{t+1}\|^2. \tag{14.38}$$

By Equation (14.37), the right-hand side of Equation (14.38) is

$$\left\| \frac{t}{t+1}(y^t - \overline{g}^t) + \frac{1}{t+1}(y^t - g^{t+1}) \right\|^2. \tag{14.39}$$

Since $d^t = \|y^t - \overline{g}^t\|$ and $\|y^t - g^{t+1}\| \leq 2M$, using Equations (14.38) and (14.39), this implies that

$$(d^{t+1})^2 \leq \left(\frac{t}{t+1}\right)^2 (d^t)^2 + \frac{4M^2}{(t+1)^2} + \frac{2t}{(t+1)^2}\langle y^t - g^{t+1}, y^t - \overline{g}^t\rangle. \tag{14.40}$$

Taking conditional expectation in both sides of Equation (14.40), conditioned on the history h^t up to stage t yields

$$\begin{aligned}&\mathbf{E}_{\sigma_1^*,\sigma_2}[(d^{t+1})^2 \mid h^t]\\ &\leq \left(\frac{t}{t+1}\right)^2 \mathbf{E}_{\sigma_1^*,\sigma_2}[(d^t)^2 \mid h^t] + \frac{4M^2}{(t+1)^2} + \frac{2t}{(t+1)^2}\mathbf{E}_{\sigma_1^*,\sigma_2}[\langle y^t - g^{t+1}, y^t - \overline{g}^t\rangle \mid h^t].\end{aligned} \tag{14.41}$$

We now show that the third element on the right-hand side of Equation (14.41) is nonpositive. If $\overline{g}^t \in C$, then $y^t = \overline{g}^t$, in which case the third element equals 0. If $\overline{g}^t \notin C$, then, because C is a B-set for Player 1, it follows from the definition of p^t that $R_1(p^t) \subset H^+(y^t - \overline{g}^t, \langle y^t - \overline{g}^t, y^t\rangle)$. Since in stage $t+1$, Player 1 plays mixed action p^t, the expected payoff in stage $t+1$, which is $\mathbf{E}_{\sigma_1^*,\sigma_2}[g^{t+1}|h^t]$, is located in $R_1(p^t)$, and therefore in $H^+(y^t - \overline{g}^t, \langle y^t - \overline{g}^t, y^t\rangle)$. It follows that

$$\langle y^t - \overline{g}^t, \mathbf{E}_{\sigma_1^*,\sigma_2}[g^{t+1} \mid h^t]\rangle \geq \langle y^t - \overline{g}^t, y^t\rangle. \tag{14.42}$$

Since the inner product is symmetric and bilinear, and since the average payoff $\overline{g}^t$ and the point y^t are determined given the history h^t, we get

$$\begin{aligned}&\mathbf{E}_{\sigma_1^*,\sigma_2}[\langle\, y^t - g^{t+1}, y^t - \overline{g}^t\rangle \mid h^t]\\ &\quad= \langle y^t - \mathbf{E}_{\sigma_1^*,\sigma_2}[g^{t+1} \mid h^t], y^t - \overline{g}^t\rangle\\ &\quad= \langle y^t - \mathbf{E}_{\sigma_1^*,\sigma_2}[g^{t+1} \mid h^t], y^t\rangle - \langle y^t - \mathbf{E}_{\sigma_1^*,\sigma_2}[g^{t+1} \mid h^t], \overline{g}^t\rangle \leq 0.\end{aligned} \tag{14.43}$$

Since the third element on the right-hand side of Equation (14.41) is nonpositive, we get

$$\mathbf{E}_{\sigma_1^*,\sigma_2}[(d^{t+1})^2 \mid h^t] \leq \left(\frac{t}{t+1}\right)^2 \mathbf{E}_{\sigma_1^*,\sigma_2}[(d^t)^2] + \frac{4M^2}{(t+1)^2}. \tag{14.44}$$

Taking the expectation over h^t of the conditional expectation on the left-hand side yields

$$\mathbf{E}_{\sigma_1^*,\sigma_2}[(d^{t+1})^2] = \mathbf{E}_{\sigma_1^*,\sigma_2}\left[\mathbf{E}_{\sigma_1^*,\sigma_2}[(d^{t+1})^2 \mid h^t]\right] \leq \left(\frac{t}{t+1}\right)^2 \mathbf{E}_{\sigma_1^*,\sigma_2}[(d^t)^2] + \frac{4M^2}{(t+1)^2}. \tag{14.45}$$

By the inductive hypothesis, $\mathbf{E}_{\sigma_1^*,\sigma_2}[(d^t)^2] \leq \frac{4M^2}{t}$, and therefore

$$\mathbf{E}_{\sigma_1^*,\sigma_2}[(d^{t+1})^2] \leq \left(\frac{t}{t+1}\right)^2 \frac{4M^2}{t} + \frac{4M^2}{(t+1)^2} = \frac{4M^2}{t+1}, \tag{14.46}$$

which is what we wanted to show. □

Recall that Markov's inequality states that for every nonnegative random variable X, and for every $c > 0$,

$$\mathbf{P}(X \geq c) \leq \frac{\mathbf{E}(X)}{c}. \tag{14.47}$$

By Lemma 14.15, and the Markov inequality (with $c = \frac{2M}{\sqrt{t}}$), we deduce that the probability that d^t is large is small (for large t):

Corollary 14.16 *For every strategy* σ_2 *of Player 2,*

$$\mathbf{P}_{\sigma_1^*,\sigma_2}\left((d^t)^2 \geq \frac{2M}{\sqrt{t}}\right) \leq \frac{2M}{\sqrt{t}}, \tag{14.48}$$

and therefore

$$\mathbf{P}_{\sigma_1^*,\sigma_2}\left(d^t \geq \frac{\sqrt{2M}}{t^{1/4}}\right) \leq \frac{2M}{\sqrt{t}}. \tag{14.49}$$

This corollary relates to the distance between $\overline{g}^t$ and the set C in stage t. We are interested in showing that this distance is small for large t, i.e., that there exists T sufficiently large such that from stage T onwards, the distance d^t remains small. In other words, while in Lemma 14.15 we show that the expected value of the random variables $(d^t)_{t\in\mathbb{N}}$ converges to 0, and therefore the sequence $(d^t)_{t\in\mathbb{N}}$ converges in probability to 0, we now wish to show that convergence occurs almost surely. Although this can be proved using the strong law of large numbers for uncorrelated random variables, we will present a direct proof of convergence, without appealing to the law of large numbers.

Lemma 14.17 *For every* $\varepsilon > 0$*, there exists a number* T *sufficiently large such that for every strategy* σ_2 *of Player 2,*

$$\mathbf{P}_{\sigma_1^*,\sigma_2}(d^t < \varepsilon, \ \ \forall t \geq T) > 1 - \varepsilon. \tag{14.50}$$

In particular, this implies that the set C is approachable by Player 1. Therefore, proving Lemma 14.17 will complete the proof of Theorem 14.14.

Proof: Let $\varepsilon > 0$. By[5] Equation (14.49), for $t = l^3$,

$$\mathbf{P}_{\sigma_1^*,\sigma_2}\left(d^{l^3} \geq \frac{\sqrt{2M}}{l^{3/4}}\right) \leq \frac{2M}{l^{3/2}}. \tag{14.51}$$

Let $L \in \mathbb{N}$. Summing Equation (14.51) over $l \geq L$ yields

$$\mathbf{P}_{\sigma_1^*,\sigma_2}\left(d^{l^3} \geq \frac{\sqrt{2M}}{l^{3/4}} \text{ for some } l \geq L\right) \leq 2M\sum_{l=L}^{\infty}\frac{1}{l^{3/2}}. \tag{14.52}$$

Consider the complement of the event on the left-hand side in Equation (14.52):

$$\mathbf{P}_{\sigma_1^*,\sigma_2}\left(d^{l^3} < \frac{\sqrt{2M}}{l^{3/4}}, \quad \forall l \geq L\right) \geq 1 - 2M\sum_{l=L}^{\infty}\frac{1}{l^{3/2}}. \tag{14.53}$$

Since the series $\sum_{l=1}^{\infty}\frac{1}{l^{3/2}}$ converges, there exists L_0 sufficiently large for $1 - 2M\sum_{l=L_0}^{\infty}\frac{1}{l^{3/2}} \geq 1 - \varepsilon$. For the remainder of the proof, we will also require that $L_0 \geq 7$.

We next prove the following lemma.

Lemma 14.18 *If* $d^{l^3} < \frac{\sqrt{2M}}{l^{3/4}}$ *for every* $l \geq L_0$*, then* $d^t < \frac{19M\sqrt{m}}{t^{1/4}}$ *for every* $t \geq (L_0)^3$.

---

[5] As will shortly be clear, the reason for setting $t = l^3$ is to ensure that the bound on the right-hand side of Equation (14.51) is a convergent series.

We will first show that Lemma 14.17 follows from Lemma 14.18. From Lemma 14.18, and Equation (14.53), one has

$$\mathbf{P}_{\sigma_1^*,\sigma_2}\left(d^t < \frac{19M\sqrt{m}}{t^{1/4}}, \quad \forall t \geq (L_0)^3\right) \geq 1 - \varepsilon, \tag{14.54}$$

from which Lemma 14.17 follows (what is the T that should be used in Lemma 14.17?).

We next turn to the proof of Lemma 14.18. Let $t \geq (L_0)^3$, and let $l \geq L_0$ be the only integer satisfying

$$l^3 \leq t < (l+1)^3. \tag{14.55}$$

We start by proving two inequalities that will be needed later.

Fact 1: $\frac{1}{l^{3/4}} \leq \frac{2}{t^{1/4}}$ for all $l \geq L_0$.
Since $l \geq 7$, one has $\left(\frac{l+1}{l}\right)^3 \leq \left(\frac{8}{7}\right)^3 < 2$, so it follows from Equation (14.55) that $\frac{t}{2} < \frac{(l+1)^3}{2} < l^3$. Therefore,

$$\frac{1}{l^{3/4}} = \frac{1}{(l^3)^{1/4}} < \frac{1}{(\frac{t}{2})^{1/4}} < \frac{2}{t^{1/4}}. \tag{14.56}$$

Fact 2: $\frac{t-l^3}{t} \leq \frac{8}{t^{1/3}}$.
Based on the definition of l (Equation (14.55)),

$$\frac{t-l^3}{t} \leq \frac{(l+1)^3 - l^3}{l^3} = \frac{3l^2+3l+1}{l^3} \leq \frac{7l^2}{l^3} = \frac{7}{l} \leq \frac{8}{l+1} < \frac{8}{t^{1/3}}, \tag{14.57}$$

where the inequality $\frac{7}{l} \leq \frac{8}{l+1}$ holds because $l \geq 7$.

Finally, the average payoff up to stage t satisfies

$$\overline{g}^t = \frac{1}{t}\sum_{n=1}^{t} g^n = \frac{1}{t}\sum_{n=1}^{l^3} g^n + \frac{1}{t}\sum_{n=l^3+1}^{t} g^n = \frac{l^3}{t}\overline{g}^{l^3} + \frac{1}{t}\sum_{n=l^3+1}^{t} g^n. \tag{14.58}$$

Since $d(x+y, x+z) = d(y,z)$ (Equation (14.10) on page 571), one has

$$\begin{aligned} d(\overline{g}^t, \overline{g}^{l^3}) &= d\left(\frac{l^3}{t}\overline{g}^{l^3} + \frac{1}{t}\sum_{n=l^3+1}^{t} g^n, \overline{g}^{l^3}\right) = d\left(\frac{1}{t}\sum_{n=l^3+1}^{t} g^n, \frac{t-l^3}{t}\overline{g}^{l^3}\right) \\ &= d\left(\frac{1}{t}\sum_{n=l^3+1}^{t} g^n, \frac{1}{t}\sum_{n=l^3+1}^{t} \overline{g}^{l^3}\right). \end{aligned} \tag{14.59}$$

Using properties of the distance relation (Equations (14.11)–(14.13) on page 571) one has

$$d(\overline{g}^t, \overline{g}^{l^3}) \leq \frac{1}{t}\sum_{n=l^3+1}^{t} d(g^n, \overline{g}^{l^3}) \leq 2M\sqrt{m}\,\frac{t-l^3}{t}. \tag{14.60}$$

Therefore,

$$
\begin{aligned}
d^t &= d(\overline{g}^t, C) && (14.61)\\
&\leq d(\overline{g}^t, \overline{g}^{l^3}) + d(\overline{g}^{l^3}, C) && (14.62)\\
&\leq 2M\sqrt{m}\,\frac{t-l^3}{t} + d^{l^3} && (14.63)\\
&\leq 2M\sqrt{m}\,\frac{t-l^3}{t} + \frac{\sqrt{2M}}{l^{3/4}} && (14.64)\\
&\leq \frac{16M\sqrt{m}}{t^{1/3}} + \frac{2\sqrt{2M}}{t^{1/4}} \leq \frac{19M\sqrt{m}}{t^{1/4}}. && (14.65)
\end{aligned}
$$

Equation (14.62) follows from triangle inequality (Equation (14.7)), Equation (14.63) follows from (14.60), Equation (14.64) follows from the assumption that $d^{l^3} \leq \frac{\sqrt{2M}}{l^{3/4}}$, and Equation (14.65) follows from Facts 1 and 2. This completes the proof of Lemma 14.18, and with it the proof of Lemma 14.17. □

Conclusion of the proof of Blackwell's Theorem (Theorem 14.14): Lemma 14.17 implies that the strategy σ_1^* guarantees with probability 1 that the distance between the average payoff $\overline{g}^t$ and the set C converges to 0 with probability 1, and therefore C is an approachable set by Player 1, which is what we wanted to prove. □

14.6 Characterizations of convex approachable sets

In the various applications making use of approachable sets, the target set is convex (we will consider two such applications later in this chapter). In this section, we will show that convex approachable sets have several simple characterizations.

Since for every mixed action p of Player 1, the set $R_1(p)$ is convex, there is a unique point $y \in R_1(p)$ that is closest to x. By the Separating Hyperplane Theorem (Theorem 23.39 on page 944), for every $x \in F \setminus R_1(p)$, the hyperplane $H(x-y, \langle x-y, y\rangle)$ separates x from $R_1(p)$, where y is the point in $R_1(p)$ closest to x. It follows that $R_1(p)$ is a B-set for Player 1, and Blackwell's Theorem (Theorem 14.14 on page 579) implies that it is an approachable set by Player 1. Moreover, the strategy under which Player 1 plays the mixed action p in every stage approaches $R_1(p)$. A similar result holds for sets $R_2(q)$ and Player 2. This leads to the following theorem:

Theorem 14.19 *For every $p \in \Delta(\mathcal{I})$, the strategy of Player 1 in which he plays the mixed action p in every stage approaches the set $R_1(p)$. For every $q \in \Delta(\mathcal{J})$, the strategy of Player 2 in which he plays the mixed action q in every stage approaches the set $R_2(q)$.*

Theorem 14.19 has the following corollary.

Corollary 14.20 *Let C be a closed set. If there exists $p \in \Delta(\mathcal{I})$ such that $C \cap R_1(p) = \emptyset$, then the set C is excludable by Player 1. If there exists $q \in \Delta(\mathcal{I})$ such that $C \cap R_2(q) = \emptyset$, then the set C is excludable by Player 2.*

The following theorem is a result in the theory of convex sets.

Theorem 14.21 *Let* H^+ *be a half-space. If for every* $p \in \Delta(\mathcal{I})$ *the set* $R_1(p)$ *is not contained in* H^+*, then there exists* $q \in \Delta(\mathcal{J})$ *such that* $R_2(q) \cap H^+ = \emptyset$.

The proof we present here applies von Neumann's Minmax Theorem (Theorem 5.11 on page 151).

Proof: Let $H^+ = H^+(\alpha, \beta) = \{x \in \mathbb{R}^m : \langle \alpha, x \rangle \geq \beta\}$. Define a two-player zero-sum game in strategic form G with the set of players $\{1, 2\}$, such that the set of pure strategies of Player 1 is $\mathcal{I}$, the set of pure strategies of Player 2 is $\mathcal{J}$, and the payoff function (for Player 1) is given by $w(i, j) = \langle \alpha, u(i, j) \rangle$. Denote by $W(p, q)$ the multilinear extension of w:

$$W(p, q) = \sum_{i=1}^{I} \sum_{j=1}^{J} p_i q_j w(i, j) = \langle \alpha, U(p, q) \rangle, \tag{14.66}$$

where U is the bilinear extension of u. By von Neumann's Theorem (Theorem 5.11 on page 151) this game has a value v. By assumption, for every mixed strategy $p \in \Delta(\mathcal{I})$ of Player 1 in the game G, there exists a mixed strategy $q \in \Delta(\mathcal{J})$ of Player 2 in G such that $U(p, q) \notin H^+$, and therefore $W(p, q)$, the payoff in G, satisfies

$$W(p, q) = \langle \alpha, U(p, q) \rangle < \beta. \tag{14.67}$$

It follows that the value v of G is less than β. Let q^* be an optimal strategy of Player 2 in G. Then $W(p, q^*) \leq v < \beta$ for every $p \in \Delta(\mathcal{I})$. It follows that $R_2(q^*) \cap H^+ = \emptyset$, as required. □

The next theorem presents conditions guaranteeing the approachability of a half-space.

Theorem 14.22 *Let* H^+ *be a half-space, and let* $p^* \in \Delta(\mathcal{I})$ *be a mixed action of Player 1. The following conditions are equivalent:*

- *(a)* $R_1(p^*) \subseteq H^+$.
- *(b)* H^+ *is a* B*-set for Player 1, and for every* $x \in F \setminus H^+$*, the mixed action* $p(x, H) = p^*$ *satisfies condition (2) of Definition 14.12 (page 579).*
- *(c) The strategy of Player 1 that plays the mixed action* p^* *in every stage approaches* H^+ *for Player 1.*

Proof: Let $H^+ = H^+(\alpha, \beta) = \{x \in \mathbb{R}^m : \langle \alpha, x \rangle \geq \beta\}$. We start by proving that (a) implies (b).[6] Let $x \in F \setminus H^+$, and let $y \in H^+$ be the point in H^+ closest to x. We will show that conditions (1) and (2) in Definition 14.12 (page 579) hold with $p(x, H^+) = p^*$. Since x is a point in $F \setminus H^+$, it will follow that H^+ is a B-set for Player 1. Since $x \notin H^+(\alpha, \beta)$, the hyperplane $H(\alpha, \beta)$ separates x and $H^+(\alpha, \beta)$, and it contains the point y. It follows (see Remark 14.13) that $H(\alpha, \beta) = H(y - x, \langle y - x, y \rangle)$. Since condition (a) holds,

---

[6] Another way to prove this is as follows: as previously stated, the set $R_1(p^*)$ is convex, and a B-set for Player 1. The half-space H^+ is a convex set containing $R_1(p^*)$, and therefore it is also a B-set (Exercise 14.11). Note that by Theorem 14.19, (a) implies (c).

$R_1(p^*) \subseteq H^+(\alpha, \beta) = H(y - x, \langle y - x, y \rangle)$. Because this is true for all $x \in F \setminus H^+$, we deduce that H^+ is a B-set for Player 1.

We next prove that (b) implies (c). Suppose, then, that condition (b) holds. By Blackwell's Theorem (Theorem 14.14 on page 579), H^+ is an approachable set. The strategy that we constructed in the proof of Blackwell's Theorem, under which Player 1 guarantees that the average payoff approaches H^+, is the strategy in which in every stage t he plays the mixed action $p(\overline{g}_{t-1}, H^+)$, where $\overline{g}_{t-1}$ is the average payoff up to stage t. Since $p(x, H) = p^*$ for every $x \in F \setminus H^+$, the strategy σ^* approaches H^+, and therefore condition (c) holds.

Finally, we prove that (c) implies (a). Let $q \in \Delta(\mathcal{J})$ be a mixed action of Player 2, and let τ be the strategy that plays the mixed action q in every stage. By the strong law of large numbers, the average payoffs under (σ^*, τ) converge to $U(p^*, q)$ with probability 1. Since σ^* approaches H^+, it follows that $U(p^*, q) \in H^+$. Since $R_1(p^*) = \{U(p^*, q) \colon q \in \Delta(\mathcal{J})\}$, and since $U(p^*, q) \in H^+$ for every $q \in \Delta(J)$, we deduce that $R_1(p^*) \subseteq H^+$, i.e., condition (a) holds. □

A *stationary strategy* of player k is a strategy that plays the same mixed action in every stage, regardless of past choices of the two players.

Corollary 14.23 *A half-space H^+ is approachable by a player if and only if the player has a stationary strategy that approaches the set.*

Proof: We prove the result for Player 1. Clearly, if Player 1 has a stationary strategy that approaches H^+, then H^+ is approachable by him. We need to prove, therefore, that if H^+ is approachable by Player 1, then Player 1 has a stationary strategy that approaches H^+. Suppose, by contradiction, that Player 1 has no stationary strategy approaching H^+. By Theorem 14.22, $R_1(p) \not\subseteq H^+$ for every $p \in \Delta(\mathcal{I})$. By Theorem 14.21, there exists a $q \in \Delta(\mathcal{J})$ such that $R_2(q) \cap H = \emptyset$. By Corollary 14.20, H is excludable by Player 2. Since a set cannot be both approachable by Player 1 and excludable by Player 2 (Theorem 14.8 on page 575), we deduce that H^+ is not approachable by Player 1, contradicting the assumption. This contradiction implies that Player 1 has a stationary strategy guaranteeing that the average payoff approaches H^+, which is what we wanted to show. □

The next theorem is, in a sense, a generalization of Theorem 14.22 to convex sets.

Theorem 14.24 *For every closed and convex set $C \subseteq \mathbb{R}^m$, the following conditions are equivalent:*

(a) For every half-space H^+ containing C, there exists $p \in \Delta(\mathcal{I})$ satisfying $R_1(p) \subseteq H^+$.
(b) The set C is a B-set for Player 1.
(c) The set C is approachable by Player 1.

Proof: We start by proving that (a) implies (b). Let $x \in F \setminus C$. Since C is closed and convex, there exists a unique point y in C closest to x. The Separating Hyperplane Theorem (Theorem 23.39 on page 944), implies that the hyperplane $H^+ := H^+(y - x, \langle y - x, y \rangle)$ separates x from C. In particular, the half-space H^+ contains C, and by (a), there exists $p \in \Delta(\mathcal{I})$ such that $R_1(p) \subseteq H^+$. We deduce from this that for every $x \in F \setminus C$ there exists $p \in \Delta(\mathcal{I})$ such that the hyperplane $H(y-x, \langle y-x, y\rangle)$ separates x from $R_1(p)$, and therefore C is a B-set for Player 1.

By Blackwell's Theorem (Theorem 14.14 on page 579), (b) implies (c).

Finally, we prove that (c) implies (a). Since C is approachable by Player 1, and since every set containing an approachable set is also approachable (Theorem 14.7 on page 575), it follows that every half-space H^+ containing C is approachable by Player 1. Corollary 14.23, and Theorem 14.22, imply that for every half-space H^+ containing C there exists $p \in \Delta(\mathcal{I})$ satisfying $R_1(p) \subseteq H^+$, and therefore condition (a) holds. □

By Theorem 14.8 (page 575), a set approachable by one player is not excludable by the other player. For convex sets, the converse statement also obtains.

Theorem 14.25 *A closed and convex set that is not approachable by one player is excludable by the other player.*

Proof: Let C be a convex and closed set that is not approachable by one of the players, say Player 1. By Theorem 14.24 (the negation of condition (a)), there exists a half-space H^+ containing C, such that for every $p \in \Delta(\mathcal{I})$, one has $R_1(p) \not\subseteq H^+$. By Theorem 14.21, there exists $q \in \Delta(\mathcal{J})$ such that $R_2(q) \cap H^+ = \emptyset$. Since H^+ contains C, it follows in particular that $R_2(q) \cap C = \emptyset$. By Corollary 14.20, it follows that C is excludable by Player 2. □

Theorem 14.25 holds only for convex sets. Example 14.4 below presents a set that is not convex, and is neither approachable by Player 1, nor excludable by Player 2.

For every vector $\alpha \in \mathbb{R}^m$, let G_α be the two-player zero-sum game (with real values payoffs) in which Player 1's set of pure strategies is $\mathcal{I}$, Player 2's set of pure strategies is $\mathcal{J}$, and the payoff function is

$$U[\alpha]_{i,j} = \langle \alpha, u(i,j)\rangle. \tag{14.68}$$

In words, using the linear transformation given by the vector α, we convert the vector payoff into a real-valued payoff. Denote the value of the game G_α by $\mathrm{val}(G_\alpha)$.

The following corollary presents a relatively simple criterion for checking whether a convex and compact set is approachable by Player 1.

Corollary 14.26 *A compact and convex set C is approachable by Player 1 if and only if*

$$\mathrm{val}(G_\alpha) \geq \min_{x \in C}\langle \alpha, x\rangle, \quad \forall \alpha \in \mathbb{R}^m. \tag{14.69}$$

Proof: We first prove that there exists a mixed action $p \in \Delta(\mathcal{I})$ such that $R_1(p) \subseteq H^+(\alpha, \beta)$ if and only if $\mathrm{val}(G_\alpha) \geq \beta$. The property $R_1(p) \subseteq H^+(\alpha, \beta)$ holds if and only if $U(p,q) \in H^+(\alpha,\beta)$ for all $q \in \Delta(\mathcal{J})$, that is, if and only if $U_\alpha(p,q) = \langle \alpha, U(p,q)\rangle \geq \beta$. In other words, the mixed action p guarantees that the payoff of the game G_α is at least β. The existence of such a mixed action p is therefore equivalent to $\mathrm{val}(G_\alpha) \geq \beta$.

For each $\alpha \in \mathbb{R}^m$ define

$$\beta_\alpha := \min_{x \in C}\langle \alpha, x\rangle. \tag{14.70}$$

The minimum is attained because C is a compact set. It follows by definition that $C \subseteq H^+(\alpha, \beta_\alpha)$ for all $\alpha \in \mathbb{R}^m$. In addition, $C \subseteq H^+(\alpha, \beta)$ if and only if $\beta \leq \beta_\alpha$.

In conclusion, C is an approachable set for Player 1 if and only if for every half-space $H^+(\alpha, \beta)$ containing it there exists $p \in \Delta(\mathcal{I})$ satisfying $R_1(p) \subseteq H^+(\alpha, \beta)$ (Theorem 14.24). Since $C \subseteq H^+(\alpha, \beta)$ if and only if $\beta \le \beta_\alpha$, this holds if and only if for all $\alpha \in \mathbb{R}^m$ there exists $p \in \Delta(\mathcal{I})$ satisfying $R_1(p) \subseteq H^+(\alpha, \beta)$ for all $\beta \le \beta_\alpha$. Since $H^+(\alpha, \beta) \supseteq H^+(\alpha, \beta_\alpha)$ when $\beta \le \beta_\alpha$, this holds if and only if for all $\alpha \in \mathbb{R}^m$ there exists $p \in \Delta(\mathcal{I})$ satisfying $R_1(p) \subseteq H^+(\alpha, \beta_\alpha)$, and this, as we have shown, is equivalent to $\mathrm{val}(G_\alpha) \ge \beta_\alpha = \min_{x \in C} \langle \alpha, x \rangle$. □

Remark 14.27 *The proof of Corollary 14.26 relies on the compactness of the set C to ensure that the minimum on the right-hand side of Equation (14.69) is attained. If the set C is not compact, yet the minimum on the right-hand side of Equation (14.69) is attained for every $\alpha \in \mathbb{R}^m$, Corollary 14.26 holds for the set C. If there exists $\alpha \in \mathbb{R}^m$ for which the minimum on the right-hand side of Equation (14.69) is not attained, Corollary 14.26 holds if we replace the minimum by an infimum (Exercise 14.21).* ◆

As we previously mentioned, Theorem 14.25, which states that every compact and convex set is either approachable by Player 1 or excludable by Player 2, does not hold for a set that is not convex. Indeed, we will now present an example of a nonconvex set C that is neither approachable by Player 1, nor excludable by Player 2.

Example 14.14 (*Continued*) The vector payoffs in this example are given by the matrix in Figure 14.9.

		Player 2	
		L	R
Player 1	T	(0, 0)	(0, 0)
	B	(1, 1)	(1, 0)

Figure 14.9 The game in Example 14.4

Define sets C_a, C_b, and C as follows (see Figure 14.10):

$$C_a := \left\{\left(\tfrac{1}{2}, y\right) : 0 \le y \le \tfrac{1}{4}\right\}, C_b := \left\{(1, y) \colon \tfrac{1}{4} \le y \le 1\right\}, \tag{14.71}$$

and

$$C := C_a \cup C_b. \tag{14.72}$$

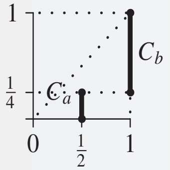


Figure 14.10 The sets C_a and C_b

We first show that Player 2 can prevent the average payoff from remaining near C, implying that C is not approachable by Player 1. Indeed, Player 2 can achieve this using the following strategy:

- If the average payoff is close to C_b, and in particular $\overline{g}_1^t$ is close to 1 (for example $\overline{g}_1^t \geq 0.9$), Player 2 plays the action R a sufficiently large number of times to make the value of the second coordinate fall (approaching 0 as t increases), and therefore the average payoff moves away from C_b.
- If the average payoff is close to C_a, and in particular $\overline{g}_1^t$ is close to $\frac{1}{2}$ (for example $0.4 \leq \overline{g}_1^t \leq 0.6$), Player 2 plays L a sufficiently large number of times to make the average payoff move towards the diagonal $x_1 = x_2$, and thus move away from C_a.
- If the average payoff is far from C (for example $0.6 < \overline{g}_1^t < 0.9$, or $\overline{g}_1^t < 0.4$), Player 2 can play any action.

We will now show that for every t, Player 1 can guarantee that in stage $2t$ the average payoff will be in C, leading to the conclusion that the set C is not excludable by Player 2.

- In the first t stages, Player 1 plays B. In particular,

$$\overline{g}^t \in \{(1, x) \colon 0 \leq x \leq 1\}. \tag{14.73}$$

- If $\overline{g}_2^t \geq \frac{1}{2}$, Player 1 plays B in the next t stages. In particular, $\overline{g}_1^{2t} = 1$, and $\overline{g}_2^{2t} \geq \frac{1}{2}\overline{g}_2^t \geq \frac{1}{4}$, and therefore $\overline{g}^{2t} \in C_b \subset C$.
- If $\overline{g}_2^t < \frac{1}{2}$, Player 1 plays T in the next t stages. In particular, $\overline{g}^{2t} = \frac{1}{2}\overline{g}^t$, i.e., $\overline{g}_1^{2t} = \frac{1}{2}$ and $\overline{g}_2^{2t} = \frac{1}{2}\overline{g}_2^t < \frac{1}{4}$. It follows that $\overline{g}^{2t} \in C_a \subset C$.

In conclusion, we have proved that C is not approachable by Player 1, and not excludable by Player 2. ◀

14.7 Application 1: Repeated games with incomplete information

Repeated games with incomplete information were first investigated by Aumann and Maschler in 1967. As the name implies, these games combine the model of repeated games (Chapter 13) with the model of games with incomplete information (Chapter 9). In such games, the payoff matrix is chosen randomly at the start of the game, and the players have different information regarding which matrix was chosen. As in all games with incomplete information, in choosing his action a player needs to take into account the state of knowledge of the other players, because a player's actions may reveal to the others some of the information he has. Because the game is repeated, the process of information revelation and information gathering becomes part of the strategic considerations of the players. In such games, the use of information means choosing an action based on the information that a player has regarding the payoff matrix. If the players know the actions taken by other players in earlier stages, the use of information may reveal information to the other players. A natural question that arises here is whether a player should use the information he has (and thus reveal it to the others), not use it, or make only partial use of it. All three cases are possible: in this section we will see an example in which it is not to a player's advantage to reveal the information he has. Exercise 14.28 presents an example in which it is to a player's advantage to reveal all of the information he has, and

Player 2

Player 1		L	R
	T	1	0
	B	0	0

The state game $s = 1$

Player 2

Player 1		L	R
	T	0	0
	B	0	1

The state game $s = 2$

Figure 14.11 The two payoff matrices of the example

Exercise 14.30 presents an example in which it is to a player's advantage to reveal only part of the information he has.

In this section we will not present a full methodological development of the subject. We will content ourselves instead with one example that exemplifies the use of repeated games with vector payoffs in the analysis of repeated games with incomplete information. The example is from Aumann and Maschler [1995]; the interested reader is encouraged to read this book for an introduction to the subject. A guided proof of the characterization of the value of two-player zero-sum repeated games with incomplete information for one player appears in Exercise 14.31.

Two players 1 and 2 play a zero-sum repeated game G, where each player has two actions: Player 1's set of actions is $\mathcal{I} = \{T, B\}$, and Player 2's set of actions is $\mathcal{J} = \{L, R\}$. As in the model of games with incomplete information, the payoff here depends on the state of nature s, which can take one of two values, $s = 1$ or $s = 2$. Each state of nature is a state game. The two state games appear in Figure 14.11. The state of nature is chosen by a fair coin toss at the beginning of the game: the probability that the chosen state is $s = 1$ is $\frac{1}{2}$. Player 1 (the row player) is informed which state of nature is chosen, but Player 2 (the column player) does not have this information. After the state of nature is chosen, and Player 1 is informed of the chosen state of nature, the two players play the infinitely repeated game whose payoff function corresponds to the chosen state of nature. At each stage t, the players know the actions chosen in all previous stages, but do not know what payoffs they have received. Player 1, however, can determine what the payoffs have been, because he knows both the actions chosen and the payoff matrix. This description of the situation is common knowledge among the players. Since Player 1 knows the true state of nature, but Player 2 does not, this game is called a repeated game with incomplete information for Player 2.

Denote the payoff matrices in Figure 14.11 by A^s, $s = 1, 2$, and denote their elements by $(a^s_{i,j})_{i \in \mathcal{I}, j \in \mathcal{J}}$.

The game presented here is a game with perfect information (Definition 6.13 on page 231) and therefore by Theorem 6.26 (page 242) it follows that every mixed strategy has an equivalent behavior strategy and vice versa. In this section we will assume that the set of strategies of every player is his set of behavior strategies. For convenience we will denote a behavior strategy of player i by σ_i instead of b_i, and the set of behavior strategies of player i will be denoted by $\mathcal{B}_i$. The information available to Player 1 in stage t is the state of nature and the sequence of actions of the players up to stage t. It follows that Player 1's behavior strategy is the function

$$\sigma_1 : \bigcup_{t \in \mathbb{N}} (\{1, 2\} \times \mathcal{I}^{t-1} \times \mathcal{J}^{t-1}) \to \Delta(\mathcal{I}). \tag{14.74}$$

In stage t, Player 2 does not know the state of nature and therefore the information available to him in stage t is the sequence of actions up to that stage. It follows that a behavior strategy of Player 2 is the function

$$\sigma_2 : \bigcup_{t \in \mathbb{N}} (\mathcal{I}^{t-1} \times \mathcal{J}^{t-1}) \to \Delta(\mathcal{J}). \tag{14.75}$$

We next define the uniform value of the game. In Chapter 13 we saw that in infinite games payoff functions are not always defined for every pair of strategies (σ_1, σ_2). As done there, we define the concept of uniform value without defining the payoff function. The definition is analogous to the definition of the uniform equilibrium we gave in Chapter 13 (Definition 13.28 on page 549), and it holds for every repeated game with incomplete information. Denote by i^t and j^t the actions chosen by the players in stage t.

Definition 14.28 *A strategy σ_1^* of Player 1* guarantees *the real number v_1 if there exists an integer T_0 such that for every $T \geq T_0$ and every strategy $\sigma_2 \in \mathcal{B}_2$,*

$$\mathbf{E}_{\sigma_1^*, \sigma_2} \left[\frac{1}{T} \sum_{t=1}^{T} a^s_{i^t, j^t} \right] \geq v_1. \tag{14.76}$$

A strategy σ_2^* of Player 2 guarantees *the real number v_2 if there exists an integer T_0 such that for every $T \geq T_0$ and every strategy $\sigma_1 \in \mathcal{B}_1$,*

$$\mathbf{E}_{\sigma_1, \sigma_2^*} \left[\frac{1}{T} \sum_{t=1}^{T} a^s_{i^t, j^t} \right] \leq v_2. \tag{14.77}$$

A real number v is called the uniform value *of the game G if for every $\varepsilon > 0$ Player 1 has a strategy guaranteeing him $v - \varepsilon$ and Player 2 has a strategy guaranteeing him $v + \varepsilon$. A strategy σ_1^* of Player 1 guaranteeing $v - \varepsilon$ for all $\varepsilon > 0$ is called an* optimal strategy *of Player 1. An optimal strategy of Player 2 is defined similarly.*

Remark 14.29 *In Definition 14.28 we defined the uniform value when the sets of strategies of the two players are $\mathcal{B}_1$ and $\mathcal{B}_2$. This concept can be defined when the sets of strategies of the players are not necessarily $\mathcal{B}_1$ and $\mathcal{B}_2$ but subsets of them.* ◆

As we will later show, the proof that Player 2 can guarantee $v + \varepsilon$ is closely related to the fact that a particular set in a repeated game with vector payoffs corresponding to G is approachable by Player 2.

Remark 14.30 *We emphasize that the state of nature is chosen only once, at the start of the game, and does not change during the course of the play of the game.* ◆

Remark 14.31 *The assumption that the players cannot see the payoffs they receive in every stage is intended to enable us to concentrate on the information directly revealed by the actions chosen by the players, and to neutralize information about the state of nature* *obtained indirectly by way of the payoffs. In a more general model, the information given to each player after each stage may be any item of information, and may include the payoff given to the players in each stage. The formal definition of a more general model of repeated games with incomplete information, and results pertaining to such a model, can be found in Zamir [1992].* ♦

Remark 14.32 *The definition of a repeated game with incomplete information for Player 2 can be generalized to any number of states of nature, as follows. A (two-player zero-sum) repeated game with incomplete information on one side is given by a finite set S of states of nature, where each state of nature is associated with a (two-player zero-sum) state game in strategic form G_s, in each of which Player 1's action set is $\mathcal{I}$, and Player 2's action set is $\mathcal{J}$. A state of nature is chosen according to a distribution $p \in \Delta(S)$, which is common knowledge among the players. The chosen state of nature is made known to Player 1, but not to Player 2. After the state of nature has been chosen and told to Player 1, the players play the (finite or infinite) repeated game whose payoff function corresponds to the chosen state of nature. At every stage t, the players know the actions that have been chosen in the previous stages, but they do not know what payoffs they have received in those stages.* ♦

As the following theorem states, if the uniform value exists, then it equals the limit of the values of the finitely repeated games, as the repetition length grows to infinity.

Theorem 14.33 *Denote by v^T the value in behavior strategies of the T-stage game Γ_T, i.e., the two-player zero-sum game whose payoff function is*

$$\gamma^T(\sigma_1, \sigma_2) = \mathbf{E}_{\sigma_1,\sigma_2}\left[\frac{1}{T}\sum_{t=1}^{T} a^s_{i^t,j^t}\right]. \tag{14.78}$$

If the uniform value v exists, then $v = \lim_{T\to\infty} v^T$.

Note that v^T, the value of the game Γ_T, exists for every T. Indeed, since the number of pure strategies of every player in this game is finite, by the Minmax Theorem (Theorem 5.11 on page 151) this game has a value in mixed strategies.

Proof: Let v be the uniform value of the game, and let $\varepsilon > 0$. Let σ_1^* and σ_2^* be strategies for each player, respectively, and let T_0 be a natural number such that for every $T \geq T_0$, and for every pair of strategies (σ_1, σ_2), Equations (14.77) and (14.76) hold. By Equation (14.77), for[7] every $T \geq T_0$,

$$v^T = \max_{\sigma_1\in\mathcal{B}_1^T}\min_{\sigma_2\in\mathcal{B}_2^T}\gamma^T(\sigma_1, \sigma_2) \tag{14.79}$$

$$\geq \min_{\sigma_2\in\mathcal{B}_2^T}\gamma^T(\sigma_1^*, \sigma_2) \geq v - \varepsilon. \tag{14.80}$$

We similarly deduce from Equation (14.76) that $v^T \leq v + \varepsilon$ for every $T \geq T_0$. Since these inequalities hold for every $\varepsilon > 0$, one has $\lim_{T\to\infty} v^T = v$, as claimed. □

---

[7] The set of behavior strategies of player i in the game Γ_T is denoted by $\mathcal{B}_i^T$.

$$\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{\text{Player 2}} \\
 & & L & R \\ \cline{3-4}
\text{Player 1} & T & \frac{1}{2} & 0 \\ \cline{3-4}
 & B & 0 & \frac{1}{2} \\ \cline{3-4}
\end{array}$$

Figure 14.12 The matrix of average payoffs

Remark 14.34 *In Chapter 13 on repeated games, we defined the concept of uniform ε-equilibrium for finite games. By restricting this concept to two-player zero-sum games, we defined the concept of uniform value defined in this section (see Exercise 13.45 on page 565). Recall that the value of the T-stage repeated game Γ_T equals the value of the one-stage (base) game for every $T \in \mathbb{N}$ (Exercise 13.8 on page 557), and that the uniform value equals the value of the one-stage (base) game (Exercise 13.45 on page 565). In particular, Theorem 14.33 holds for repeated games with complete information. Furthermore, in repeated games with complete information the uniform value coincides with the value of the one-stage (base) game. As we will see shortly this does not necessarily hold in repeated games with incomplete information.* ♦

In the game described above, Player 1 has information that Player 2 lacks – knowledge of the chosen state of nature. What is the "best" way for him to use this information? We will first show that if Player 1 ignores the information he has regarding the chosen state of nature, he cannot guarantee more than $\frac{1}{4}$, while if Player 2 ignores the information revealed by the actions of Player 1, he cannot guarantee less than $\frac{1}{2}$.

Denote by $\mathcal{B}_1^*$ the set of behavior strategies of Player 1 in the repeated game that do not use the information he has regarding the state of nature, i.e., the mixed actions played in each stage that are independent of s (and depend only on the actions of the players in previous stages).

Proposition 14.35 The uniform value of the game with the sets of strategies $\mathcal{B}_1^*$ and $\mathcal{B}_2$ is $\frac{1}{4}$.

Proof: If Player 1 ignores his information on the state of nature, the players actually play the repeated game with the average payoff matrices appearing in Figure 14.12.

Indeed, if Player 1 does not use the information in his possession on the state of nature, the game is equivalent to the game in which the player does not know the state of nature. In that game, the expected payoff in each stage is given by the matrix appearing in Figure 14.12. The value in mixed strategies of the one-stage game, as well as the uniform value of the repeated game, is $\frac{1}{4}$ (verify!). □

As the following claim shows, if Player 2 does not use the information revealed by the actions of Player 1, he cannot guarantee less than $\frac{1}{2}$. Denote by $\mathcal{B}_2^*$ the set of strategies of Player 2 in the repeated game that are independent of the history of the actions of Player 1.

Proposition 14.36 The uniform value of the game with sets of strategies $\mathcal{B}_1^*$ and $\mathcal{B}_2$ is $\frac{1}{2}$.

Proof: To show that the uniform value of the game is $\frac{1}{2}$, we will construct strategies for the two players that guarantee this value. Consider first the following strategy σ_2^* of Player 2: in each stage, play the mixed action $[\frac{1}{2}(L), \frac{1}{2}(R)]$. For every strategy σ_1 of Player 1, the expected payoff in each stage is at most $\frac{1}{2}$, and therefore the strategy σ_2^* guarantees $\frac{1}{2}$ for Player 2.

Consider next the following strategy σ_1^* of Player 1: if $s = 1$, play T in every stage; if $s = 2$, play B in every stage. The mixed action that Player 2 plays in stage t is independent of the action played by Player 1 in the previous stages. Denote this mixed action by y^t (which may depend on the actions played by Player 2 in previous stages); this is the probability of playing L in stage t. For every strategy σ_2 of Player 2,

$$\mathbf{E}_{\sigma_1^*,\sigma_2}\left[\frac{1}{T}\sum_{t=1}^{T} a^s_{i^t,j^t}\right] = \mathbf{E}_{\sigma_1^*,\sigma_2}\left[\frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{1}_{\{s=1\}}y^t + \mathbf{1}_{\{s=2\}}(1-y^t)\right)\right] \tag{14.81}$$

$$= \mathbf{P}(s=1)\mathbf{E}_{\sigma_1^*,\sigma_2}\left[\frac{1}{T}\sum_{t=1}^{T} y^t\right] + \mathbf{P}(s=2)\mathbf{E}_{\sigma_1^*,\sigma_2}\left[\frac{1}{T}(1-y^t)\right] \tag{14.82}$$

$$= \tfrac{1}{2}\mathbf{E}_{\sigma_1^*,\sigma_2}\left[\tfrac{1}{T}\textstyle\sum_{t=1}^{T} y^t\right] + \tfrac{1}{2}\mathbf{E}_{\sigma_1^*,\sigma_2}\left[\tfrac{1}{T}(1-y^t)\right] \tag{14.83}$$

$$= \mathbf{E}_{\sigma_1^*,\sigma_2}\left[\frac{1}{T}\sum_{t=1}^{T}\frac{1}{2}\right] = \frac{1}{2}. \tag{14.84}$$

Since the strategy σ_1^* of Player 1 guarantees $\frac{1}{2}$, it guarantees $\frac{1}{2} - \varepsilon$ for all $\varepsilon > 0$. Since the strategy σ_2^* of Player 2 guarantees $\frac{1}{2}$, it guarantees $\frac{1}{2} + \varepsilon$ for all $\varepsilon > 0$. It follows that the uniform value of the game is indeed $\frac{1}{2}$. □

Can Player 1 guarantee more than $\frac{1}{4}$ by using the information he has? If the game is a one-stage game, the answer is affirmative: the strategy under which Player 1 plays T if $s = 1$ and B if $s = 2$ guarantees him $\frac{1}{2}$ (Exercise 14.26). Intuitively, it might seem that in the repeated game, if the state of nature is $s = 1$, Player 1 would want to play T more often, while if the state of nature is $s = 2$, he would want to play B more often. But in the repeated game, this strategy is not necessarily a good one, because it reveals the state of nature to Player 2: if Player 2 notices that Player 1 plays T more often, he will ascribe greater probability to the event that the payoff matrix corresponds to $s = 1$ and therefore he will increase the probability that he will play R, while if he notices that Player 1 plays B more often, he will increase the probability that he will play L. As the following proposition shows, in the long run, even when he uses the information he has, Player 1 cannot guarantee that the average payoff will be significantly higher than $\frac{1}{4}$. To prove this proposition, we use tools we developed for approachable sets.

Proposition 14.37 The uniform value of the game G (with sets of strategies $\mathcal{B}_1$ and $\mathcal{B}_2$) is $\frac{1}{4}$.

Proof: In Proposition 14.35, we saw that Player 1 can guarantee $\frac{1}{4}$. It remains to show that Player 2 can guarantee $\frac{1}{4} + \varepsilon$ for every $\varepsilon > 0$. To do so, consider the repeated game with two-dimensional vector payoffs G_V, where the sets of actions are $\mathcal{I} = \{T, B\}$ and

	L	R
T	$(1, 0)$	$(0, 0)$
B	$(0, 0)$	$(0, 1)$

Figure 14.13 The matrix of vector payoffs in the game G_V

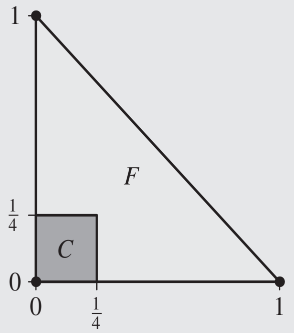


Figure 14.14 The sets F and C of the example

$\mathcal{J} = \{L, R\}$, and the two coordinates of the payoff represent the payoff in each of the two states of nature (see Figure 14.13).

In this game, the set of possible payoffs is $F = \text{conv}\{(0,0), (0,1), (1,0)\}$ (see Figure 14.14). Define a set C as follows:

$$C = \left[0, \tfrac{1}{4}\right]^2 = \left[0, \tfrac{1}{4}\right] \times \left[0, \tfrac{1}{4}\right] = \left\{(x, y) \colon 0 \leq x \leq \tfrac{1}{4}, 0 \leq y \leq \tfrac{1}{4}\right\}. \qquad (14.85)$$

The proposition will be proved by proving the following two claims:

Claim 14.38 *The set C is approachable by Player 2 in the repeated game G_V.*

Claim 14.39 *If the set C is approachable by Player 2 in the repeated game G_V, then the uniform value of the game G is $\frac{1}{4}$.*

We begin by proving the second claim.

Proof of Claim 14.39: Suppose that the set C is approachable by Player 2 in the game G_V, and denote by σ_2^* his strategy that approaches this set. Then for every $\varepsilon > 0$ there exists $T \in \mathbb{N}$ such that for every strategy σ_1 of Player 1 in the game G_V:

$$\mathbf{P}_{\sigma_1, \sigma_2^*}\left(d(\overline{g}^t, C) < \varepsilon, \quad \forall t \geq T\right) > 1 - \varepsilon. \qquad (14.86)$$

Since C is the square $[0, \frac{1}{4}]^2$, if x is close to C, the coordinates x_1 and x_2 cannot be much greater than $\frac{1}{4}$:

$$\mathbf{P}_{\sigma_1,\sigma_2^*}\left(\overline{g}_1^t < \tfrac{1}{4} + \varepsilon, \quad \forall t \geq T\right) > 1 - \varepsilon, \tag{14.87}$$

$$\mathbf{P}_{\sigma_1,\sigma_2^*}\left(\overline{g}_2^t < \tfrac{1}{4} + \varepsilon, \quad \forall t \geq T\right) > 1 - \varepsilon. \tag{14.88}$$

Since Player 2 does not know the state of nature, the strategy σ_2^* in G_V is also a strategy in the game G. Player 1, in contrast, knows the state of nature, and so his strategy σ_1 in G is essentially composed of two "sub-strategies," each of which is a strategy in G_V: the strategy σ_1^1 that he plays if the state of nature is $s = 1$, and the strategy σ_1^2 that he plays if the state of nature is $s = 2$.

Equations (14.87)–(14.88) hold for every strategy σ_1 of Player 1, and in particular for the strategies σ_1^1 and σ_1^2. Inserting the strategy σ_1^1 in Equation (14.87) and the strategy σ_1^2 in Equation (14.88) yields

$$\mathbf{P}_{\sigma_1^1,\sigma_2^*}\left(\overline{g}_1^t < \tfrac{1}{4} + \varepsilon, \quad \forall t \geq T\right) > 1 - \varepsilon, \tag{14.89}$$

$$\mathbf{P}_{\sigma_1^2,\sigma_2^*}\left(\overline{g}_2^t < \tfrac{1}{4} + \varepsilon, \quad \forall t \geq T\right) > 1 - \varepsilon. \tag{14.90}$$

Equation (14.89) states that when $s = 1$, with probability close to 1, the average payoff when Player 1 plays according to strategy σ_1^1 cannot be much more than $\frac{1}{4}$. Equation (14.90) says the same is true when $s = 2$ for the strategy σ_1^2. Equations (14.89)–(14.90) imply that for every $t \geq T$

$$\mathbf{E}_{\sigma_1,\sigma_2^*}[\overline{g}^t] \leq (1 - \varepsilon)\left(\tfrac{1}{4} + \varepsilon\right) + \varepsilon \leq \tfrac{1}{4} + 2\varepsilon. \tag{14.91}$$

It follows that the strategy σ_2 of Player 2 guarantees $\frac{1}{4} + \varepsilon$ for all $\varepsilon > 0$, and therefore the uniform value of the game is $\frac{1}{4}$. □

Proof of Claim 14.38: Recall that $R_2(q)$ is the set in which the expected payoffs are located when Player 2 plays the mixed action $[q(L), (1-q)(R)]$ in each stage (see Equation (14.24)). In the game G_V,

$$R_2(q) = \text{conv}\{(q, 0), (0, 1 - q)\}. \tag{14.92}$$

Figure 14.15 shows the set C and the sets $R_2(q)$ for five values of q.

We now check that C is a B-set for Player 2 in the game G_V. Let x be a vector payoff in $F \setminus C$. If $\frac{1}{4} \leq x_1 \leq 1 - x_2$, and $0 \leq x_2 \leq \frac{1}{4}$ (the area labeled A_1 in Figure 14.16(a)), the point in C closest to x is $y = (\frac{1}{4}, x_2)$, the hyperplane separating x from C passes through this point, and perpendicular to the line interval xy is the line $x_1 = \frac{1}{4}$, which separates x from $R_2(0)$. If $0 \leq x_1 \leq \frac{1}{4}$ and $\frac{1}{4} \leq x_2 \leq 1 - x_1$ (the area labeled A_2 in Figure 14.16(b)), the point in C closest to x is $y = (x_1, \frac{1}{4})$, the hyperplane separating x from C passes through this point, and perpendicular to the line interval xy is the line $x_2 = \frac{1}{4}$, which separates x from $R_2(1)$.

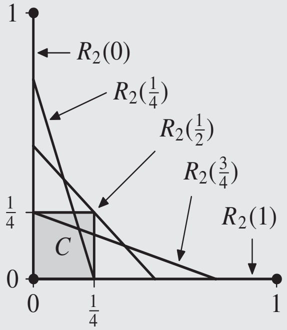


Figure 14.15 The set C and five of the sets $R_2(q)$

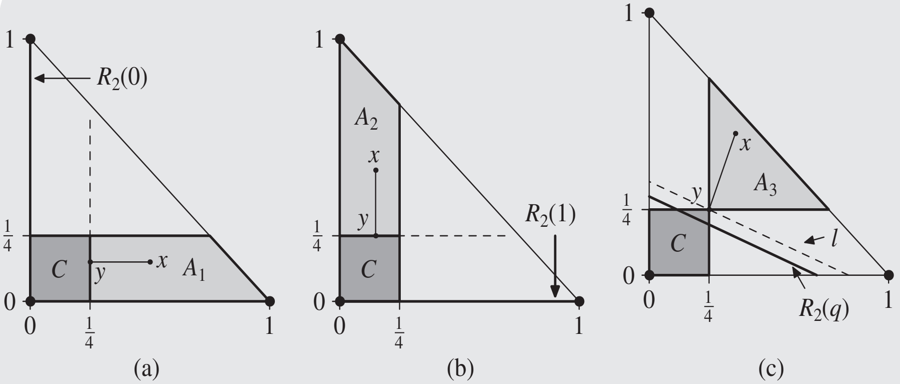


Figure 14.16 The set C is approachable by Player 2

If $x_1 \geq \frac{1}{4}$, $x_2 \geq \frac{1}{4}$, and $x_1 + x_2 \leq 1$ (the area labeled A_3 in Figure 14.16(c)), the point C closest to x is $y = (\frac{1}{4}, \frac{1}{4})$. The hyperplane l separating x from C passing through this point and perpendicular to the line interval xy can be calculated using similar triangles. In Figure 14.17, which is a more detailed view of Figure 14.16(c), we see that the triangles xyz, yab, and dyc are similar triangles. Since the lengths of the sides xz and yz are known (these lengths are $x_2 - \frac{1}{4}$ and $x_1 - \frac{1}{4}$ respectively), and since the lengths of the sides by and cy are also known (they are both $\frac{1}{4}$), one can calculate the lengths of the sides ab and cd, deducing that the hyperplane l is the line passing through the points $(0, w_2)$ and $(w_1, 0)$, where $w_2 = \frac{1}{4} + \frac{\frac{1}{4}}{x_2 - \frac{1}{4}}\left(x_1 - \frac{1}{4}\right)$ and $w_1 = \frac{1}{4} + \frac{\frac{1}{4}}{x_1 - \frac{1}{4}}\left(x_2 - \frac{1}{4}\right)$. Since $a^2 + b^2 \geq 2ab$ for every $a, b \in \mathbb{R}$,

$$w_1 + w_2 = \frac{1}{2} + \frac{1}{4} \frac{\left(x_1 - \frac{1}{4}\right)^2 + \left(x_2 - \frac{1}{4}\right)^2}{\left(x_1 - \frac{1}{4}\right)^2 \left(x_2 - \frac{1}{4}\right)^2} \geq \frac{1}{2} + \frac{1}{2} = 1. \tag{14.93}$$

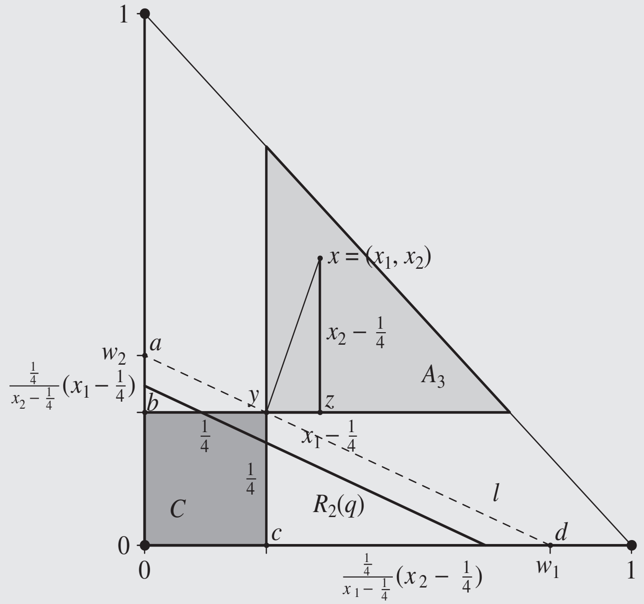


Figure 14.17 The hyperplane l separating x and $R_2(q)$

Defining

$$q = \frac{w_1}{w_1 + w_2} = \frac{x_2 - \frac{1}{4}}{x_1 + x_2 - \frac{1}{2}}, \tag{14.94}$$

we get that $R_2(q)$ is parallel to the line l (see Figure 14.16(c)); therefore the line l separates $R_2(q)$ from x. In summary, we showed that C is indeed a B-set for Player 2. □

The proof of Proposition 14.37 is complete. □

Using the proof of the last claim, we can now describe an optimal strategy for Player 2:

- If $\overline{g}^t \in C$, play any action.
- If $\overline{g}^t \in A_1$ play R.
- If $\overline{g}^t \in A_2$ play L.
- If $\overline{g}^t \in A_3$ play the mixed action $[q(L), (1-q)(R)]$, where q is defined in Equation (14.94).

In the first case, the average payoff up to that stage of the game is not greater than $\frac{1}{4}$, whether $s = 1$ or $s = 2$, and so Player 2 can play any action. In the second case, if $s = 1$, the average payoff is greater than $\frac{1}{4}$, while if $s = 2$, the payoff is not greater than $\frac{1}{4}$. It follows that Player 2 needs to guarantee that the average payoff is reduced if $s = 1$, and therefore must play R. The third case is analogous to the second case. In the fourth case, the average payoff is greater than $\frac{1}{4}$, whether $s = 1$ or $s = 2$. In this case, Player 2 plays a mixed action that guarantees that (in expectation) both coordinates go down. The greater the ratio $\frac{\overline{g}_1^t}{\overline{g}_2^t}$ is, the more the average payoff when $s = 1$ is greater than the average payoff when $s = 2$; therefore the probability that Player 2 plays R grows. Note that this strategy depends on the history only by way of the average payoff over the course of the game up to that stage.

14.8 Application 2: Challenge the expert

In this section, we present a second application of the theory developed in this chapter, this time to the dynamic process of decision making. Most of us seek the advice of experts: statesmen and government leaders employ experts in many different fields: media advisers, policy advisers, and legal advisers; investors listen to the advice of financial advisers; the person-in-the-street listens to the weather forecast to decide what to wear tomorrow, and where to go on a trip. It is sometimes the case that a decision maker has several different advisers, and needs to decide whose advice to take: government leaders often have multiple staffs working for them; banks employ many financial advisers, and on different television channels one may see weather forecasters using different models for weather prediction. A decision maker's problems are not resolved if he has many experts at his disposal – the problem has just been transformed: the problem now is not which decision to take, but which expert to heed.

In this section we will consider the problem of dynamic decision making. By this, we mean that in each stage, the decision maker needs to adopt the advice of one expert, out of a group of experts, and then choose the action recommended by that expert. In each stage the decision maker receives a payoff (= utility) that is determined by the action chosen, and the "state of nature," i.e., the environment in which the action is taken. The payoff to an investor depends both on the specific investment chosen, and on the behavior of the market; the payoff to a head of state depends on his or her actions, and on the actions of other heads of state.

We further assume that it is unknown how the state of nature changes, or, alternatively, that the way the state of nature changes is so complex that the cost or the time investment required for computing it is vast. In addition, we suppose that the actions chosen by the decision maker do not change the way the state of nature changes. This is a reasonable assumption to make with respect to a person deciding what he or she will wear based on the weather forecast, or a person deciding where to invest \$10,000, in accordance with advice received from a financial expert, but is not reasonable for a head of state, whose decision may influence the decisions of other heads of state.

Suppose that the goal of a decision maker is to be at least as good as each one of the advisers. In other words, the decision maker is interested in ensuring that his average payoff is no less than the average payoff of a person who from the start picks one of the advisers, and always listens to his or her advice. If this is possible, it means that the decision maker's performance is at least as good as that of the best expert.

14.8.1 The model

With the general presentation of the problem behind us, we turn our attention to the formal definitions.

Definition 14.40 *A* decision problem with experts *is given by:*

- *A finite set S of states of nature.*
- *A finite set A of actions.*
- *A payoff function $u : A \times S \to \mathbb{R}$.*
- *A finite set E of experts.*

The interpretation we give to this model is as follows. In every stage $t \in \mathbb{N}$, the state of nature s^t is one of the states in S, and each expert $e \in E$ recommends[8] an action $a^t_e \in A$. The decision maker, who does not know what is the true state of nature, must choose one expert from E. If the decision maker chooses expert e, and the state of nature is $s \in S$, the decision maker receives the payoff $u(a^t_e, s^t)$. We assume that the decision maker learns what the state of nature is after choosing the expert whose recommendation he followed.[9] We do not assume anything regarding the information that the experts have: it is possible that they have full or partial information about the states of nature $(s^t)_{t\in\mathbb{N}}$, they may receive information from time to time about future states of nature, and some may have full or partial information about the advice given by the other experts.

For every distribution $\alpha \in \Delta(A)$ over the set of actions, define

$$U(\alpha, s) := \sum_{a\in A} \alpha(a)u(a, s). \tag{14.95}$$

This is the expected payoff if the state of nature is s and the action that is chosen by the probability distribution is α.

The information that the decision maker has in stage t includes the states of nature in the past (in stages $1, 2, \ldots, t-1$), the expert chosen in each past stage, and the recommendations of the experts in each past stage. The set of all possible histories in stage t is

$$H(t) := (S \times E \times A^{|E|})^{t-1}. \tag{14.96}$$

Denote a history in $H(t)$ by $h^t = (s^1, e^1, a^1, \ldots, s^{t-1}, e^{t-1}, a^{t-1})$, where $a^j = (a^j_e)_{e\in E}$ is the vector composed of the recommendations of the experts in stage j.

Definition 14.41 *A decision maker's (behavior)* strategy *in a decision problem with experts is a function*

$$\sigma : \bigcup_{t=1}^{\infty} H(t) \to \Delta(E). \tag{14.97}$$

---

[8] For simplicity, we assume that each expert recommends an action, rather than a lottery over the actions. The results presented in this section hold if we assume that each expert recommends a mixed action in each stage; the only difference is that the proofs become more complex.

[9] For the analysis we present here, all we need to assume is that the decision maker knows the payoff received in each stage, and what payoff he would have received if he had followed the recommendation of each other expert.

In other words, in every stage, the decision maker needs to choose, based on his past choices and the performances of the experts in the past, one expert, and to implement the action recommended by that expert. The decision maker may choose an expert in each stage by lottery. If the decision maker adopts strategy σ, after history $h^t = (s^1, e^1, a^1, \dots, s^{t-1}, e^{t-1}, a^{t-1})$ he chooses one of the experts in E using the probability distribution $\sigma(h^t)$: $\sigma(e; h^t)$ is the probability that expert e is chosen. Since in stage t expert e recommends the action a^t_e, the average payoff that the decision maker receives up to stage T is

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E}\sigma(e;h^t)u\left(a^t_e, s^t\right) = \sum_{e\in E}\left(\frac{1}{T}\sum_{t=1}^{T}\sigma(e;h^t)u\left(a^t_e, s^t\right)\right), \tag{14.98}$$

where s^t is the state of nature in stage t, and h^t is the history up to stage t. If the decision maker would follow the recommendation of a particular expert e, his average payoff up to stage T would be

$$\frac{1}{T}\sum_{t=1}^{T}u\left(a^t_e, s^t\right). \tag{14.99}$$

We are using here the assumption that the action implemented by the decision maker does not affect the way that the state of nature changes, and therefore if $(s^t)_{t\in\mathbb{N}}$ is the sequence of states of nature that obtain when the decision maker implements strategy σ, that same sequence would obtain if he were always to choose expert e. The decision maker's goal is to attain a performance at least as good as the performance of any one of the experts.

Definition 14.42 *A decision maker's strategy σ is a* no-regret strategy *if for each expert $e \in E$, and each sequence $(s^1, s^2, \dots)$ of states of nature,*

$$\mathbf{P}_\sigma\left(\liminf_{T\to\infty}\left(\frac{1}{T}\sum_{t=1}^{T}u\left(a^t_{e^t}, s^t\right) - \frac{1}{T}\sum_{t=1}^{T}u\left(a^t_e, s^t\right)\right) \geq 0\right) = 1. \tag{14.100}$$

In words, a strategy is a no-regret strategy if, with probability 1, the decision maker does not regret the way he played: his performance is at least as good as the hypothetical performance of any expert, independently of the way the state of nature changes. If a decision maker has a no-regret strategy, he can pride himself in making decisions at least as good as each expert.

Remark 14.43 *Although the model we have presented here is a single-player model, where the single player is the decision maker, it is convenient to think of this situation as including an additional player, nature, who chooses the state of nature, and we imagine that nature seeks to reduce the decision maker's payoff as much as possible. By doing so, we have defined a two-player zero-sum repeated game.* ◆

Does a no-regret strategy exist? If so, can we construct it, or must we content ourselves with proving its theoretical existence, without finding a no-regret strategy in a practical way? Using the theory of approachability developed in this chapter, we will prove the following theorem.

Theorem 14.44 *The decision maker has a no-regret strategy.*

In addition, we will construct this strategy in a way that can easily be programmed into a computer.

14.8.2 Existence of a no-regret strategy: a special case

Recall that for every finite set X the number of elements in X is denoted by $|X|$. We first present a proof in the simple case where there are exactly $|A|$ experts. Each expert is identified with the action that he recommends in each stage. In other words, for every action $a \in A$ there is an expert e such that $a_e^t = a$ in every stage t. We will later show how the proof needs to be changed for the general case. Since the recommendations of the experts are fixed, we can define a strategy to be a function associating for each $t \in \mathbb{N}$, each sequence of length $t-1$ of actions chosen by the decision maker in the past, and each sequence of length $t-1$ of states of nature, with a probability distribution over the set of actions.

Define an auxiliary game with vector payoffs G_V as follows. In the auxiliary game, Player 1 represents the decision maker whose set of actions is A, and Player 2 represents nature, whose set of actions is S. The payoffs are $|A|$-dimensional and the payoff function w is defined as follows: the vector payoff $w(a,s) = (w_b(a,s))_{b\in A}$ that obtains when Player 1 implements action a and Player 2 implements s is given by

$$w_b(a,s) := u(a,s) - u(b,s). \tag{14.101}$$

This is the difference between the payoff received by the decision maker if he chooses action a (based on the recommendation of expert a, who always recommends action a) and the payoff received by the decision maker if he chooses action b (based on the recommendation of expert b, who always recommends action b).

A strategy for Player 1 in this auxiliary game with vector payoffs is a function

$$\widehat{\sigma} : \bigcup_{t=1}^{\infty} (A \times S)^{t-1} \to \Delta(A). \tag{14.102}$$

In words, a strategy associates every finite sequence of actions chosen by Player 1 and states of nature chosen by Player 2 with a mixed action. As previously noted, $\widehat{\sigma}$ is a strategy in the decision problem with experts, in the special case.

Denote by

$$C = \{x \in \mathbb{R}^A : x_i \geq 0 \quad \forall i \in A\} \tag{14.103}$$

the nonnegative orthant in $\mathbb{R}^A$. To prove that there exists a no-regret strategy, and to describe such a strategy in detail, we will prove the following theorems.

Theorem 14.45 *The set C is approachable by Player 1 in the auxiliary game with vector payoffs G_V.*

Theorem 14.46 *Every strategy σ of Player 1 that approaches the set C in G_V is a no-regret strategy for the decision maker in the decision problem with experts.*

We will first show that these two theorems prove Theorem 14.44. We will then proceed to prove them.

By Theorem 14.45, the set C is approachable by Player 1 in the game G_V by a strategy in the auxiliary game that approaches C. By Theorem 14.46, the strategy σ is a no-regret strategy for the decision maker in the decision problem with experts, and Theorem 14.44 is proved.

Proof of Theorem14.45: For every $\pi \in \mathbb{R}^A$ define a two-player zero-sum one-stage game $\widehat{G}_\pi$ in which the set of pure strategies of Player 1 is A, the set of pure strategies of Player 2 is S, and the payoff function $\widehat{W}_\pi$ is defined by

$$\widehat{W}_\pi = \langle \pi, w(a,s)\rangle. \tag{14.104}$$

We will show that for every $\pi \in \mathbb{R}^A$,

$$\operatorname{val}(\widehat{G}_\pi) \geq \inf_{x\in C} \langle \pi, x\rangle, \tag{14.105}$$

and Theorem 14.45 then follows from this equation and Corollary 14.26 on page 588, because the set C is convex.

If $\pi = \vec{0}$, both sides of Equation (14.105) are zero, and the inequality holds. If there exists an index a such that $\pi_a < 0$, then the right-hand side of Equation (14.105) is $-\infty$. To see this, for each $k \in \mathbb{N}$ denote by x^k the vector $(0, \ldots, 0, k, 0, \ldots, 0)$, all of whose coordinates are 0, except for coordinate a, which equals k. Since $x^k \in C$,

$$\inf_{x\in C} \langle \pi, x\rangle \leq \inf_{k\in\mathbb{N}} \langle \pi, x^k\rangle = -\infty. \tag{14.106}$$

In contrast, the left-hand side of Equation (14.105) is finite, being the value of a two-player zero-sum game. The inequality in Equation (14.105) therefore holds in this case.

It remains to check the case where all the coordinates of π are nonnegative, and at least one of them is strictly positive. Since both sides of Equation (14.105) are linear in π (i.e., $\operatorname{val}(\widehat{G}_{\lambda\pi}) = \lambda \operatorname{val}(\widehat{G}_\pi)$ and $\langle \lambda\pi, x\rangle = \lambda\langle \pi, x\rangle$ for every $\lambda > 0$), by multiplying both sides of Equation (14.105) by $\lambda = \frac{1}{\sum_{a\in A}\pi_a}$, which is a positive value, we may assume without loss of generality that $\sum_{a\in A}\pi_a = 1$, and then we may interpret π as a mixed strategy of Player 1. We will show that in this case

$$\operatorname{val}(\widehat{G}_\pi) \geq 0 \geq \inf_{x\in R} \langle \pi, x\rangle. \tag{14.107}$$

The right-hand inequality in Equation (14.107) holds, because $\vec{0} \in C$, and therefore

$$\inf_{x\in C} \langle \pi, x\rangle \leq \langle \pi, \vec{0}\rangle = 0. \tag{14.108}$$

We next turn our attention to the proof of the left-hand inequality of Equation (14.107). Denote by $\mathbf{1}_s \in \mathbb{R}^S$ the column vector in which the coordinate s equals 1, and all the other coordinates equal 0. For each $s \in S$ the vector $\mathbf{1}_s$ corresponds to the pure strategy s of

Player 2. We now show that $\widehat{W}_\pi(\pi, \mathbf{1}_s) \geq 0$ for each $s \in S$.

$$\widehat{W}_\pi(\pi, \mathbf{1}_s) = \sum_{e \in E} \pi(e) \sum_{e' \in E} \pi(e')(u(e,s) - u(e',s)) \tag{14.109}$$

$$= \sum_{e \in E} \pi(e) \left(u(e,s) - \sum_{e' \in E} \pi(e') u(e',s) \right) \tag{14.110}$$

$$= \sum_{e \in E} \pi(e) u(e,s) - \sum_{e' \in E} \pi(e') u(e',s) = 0. \tag{14.111}$$

It follows that $\widehat{W}_\pi(\pi, y) = 0$ for every mixed action $y \in \Delta(S)$. In other words, the mixed strategy π guarantees Player 1 the payoff 0; therefore $\mathrm{val}(\widehat{G}_\pi) \geq 0$, which is the left-hand inequality in Equation (14.105). □

Proof of Theorem 14.46: Let σ be a strategy of Player 1 in the auxiliary game G_V that approaches C. Then for every $\varepsilon > 0$, there exists $T_0 \in \mathbb{N}$ such that for every strategy τ of Player 2 in G_V,

$$\mathbf{P}_{\sigma,\tau} \left(d \left(\frac{1}{T} \sum_{t=1}^{T} w(a^t, s^t), C \right) \leq \varepsilon \quad \forall T \geq T_0 \right) > 1 - \varepsilon. \tag{14.112}$$

This equation implies that

$$\mathbf{P}_{\sigma,\tau} \left(\lim_{T \to \infty} d \left(\frac{1}{T} \sum_{t=1}^{T} w(a^t, s^t), C \right) = 0 \right) = 1. \tag{14.113}$$

Since $w_b(a,s) = u(a,s) - u(b,s)$, and since C is the nonnegative orthant, we deduce that for every $b \in A$,

$$\mathbf{P}_{\sigma,\tau} \left(\liminf_{T \to \infty} \left(\frac{1}{T} \sum_{t=1}^{T} u(a^t, s^t) - \frac{1}{T} \sum_{t=1}^{T} u(b, s^t) \right) \geq 0 \right) = 1. \tag{14.114}$$

This equation holds for every strategy τ of Player 2, and in particular for every sequence of states $(s^1, s^2, \ldots)$ (constituting a pure strategy) in which Player 2 plays the action s^t in stage t. This states precisely that σ is a no-regret strategy. □

14.8.3 Existence of a no-regret strategy: the general case

We now show how the proof of Theorem 14.44 needs to be changed when we do not assume that $E = A$ and that for every action $a \in A$ there exists an expert recommending that action.

Two changes are implemented in the definition of the game G_V. First of all, since the set of experts is E, the set of actions of Player 1 is E (and not A). Secondly, since the recommendations of the experts change from stage to stage, the payoff function W is also dependent on the stage t. For every $e \in E$ and $s \in S$, and every stage t, the vector $w^t(e,s)$ is in $\mathbb{R}^E$, and is defined by

$$W^t_{e'}(e,s) = u\left(a^t_e, s\right) - u\left(a^t_{e'}, s\right). \tag{14.115}$$

The rest of the proof of Theorem 14.45 is similar to the proof above, with the only difference being that now π is a vector in $\mathbb{R}^E$, and when $\pi_e \geq 0$ for all $e \in E$ and $\sum_{e \in E} \pi_e = 1$, we interpret the vector π as a mixed action of Player 1 (Exercise 14.32). The proof of Theorem 14.46 needs to be adapted to the case in which the payoffs change during the game (Exercise 14.25).

14.9 Discussion

In this chapter, we described a model of two-player repeated games with vector payoffs. The central concept of the model is that of a set approachable by a player. We gave a geometric sufficient condition for a closed set to be approachable: every closed set containing a B-set for a player is approachable by him. Lehrer [2003] generalizes Blackwell's Theorem (Theorem 14.14 on page 579) to the case in which the payoffs are in an infinite-dimensional space rather than an m-dimensional space.

Hou [1971] proves the following theorem:

Theorem 14.47 *A closed set C is approachable by player k if and only if it contains a B-set for player k.*

Another proof of the same theorem, arrived at independently, appears in Spinat [2002]. Since the set C is excludable by a player if and only if there exists $\delta > 0$ such that the set $\{x \in \mathbb{R}\colon d(x, C) \geq \delta\}$ is approachable by him, Theorems 14.14 and 14.47 provide a geometric characterization of excludable sets.

Every closed and convex set is either approachable by one of the players, or excludable by the other player (Theorem 14.25, page 588). We saw an example of a set C that is neither approachable by Player 1 nor excludable by Player 2 (this set is necessarily not convex). In that example, although the set C is not approachable by Player 1, for every t sufficiently large, Player 2 has a strategy guaranteeing that the average payoff in stage t is close to C. A set satisfying this property is called "weakly approachable" by Player 1.

Definition 14.48 *A set S is* weakly approachable *by player k if for every $\varepsilon > 0$ there exists $T \in \mathbb{N}$ such that for every $t \geq T$ there exists a strategy σ_k of player k (that depends on t) satisfying the property that for every strategy σ_{-k} of the other player*

$$\mathbf{E}_{\sigma_k, \sigma_{-k}}(d(\overline{g}^t, C)) \leq \varepsilon. \tag{14.116}$$

A set S is weakly excludable *by player k if there exists $\delta > 0$ such that the set $\{x \in \mathbb{R}^m : d(x, S) \geq \delta\}$ is weakly approachable by player k.*

The difference between the concept of "approachability" and the concept of "weak approachability" is subtle: the concept of "approachability" requires the existence of a strategy guaranteeing that the average payoff up to stage t is close to C for all $t \geq T$ (i.e., the same strategy is good for all stages $t \geq T$), while in the concept of "weak approachability" the strategy can depend on t (and it may be the case that there is no strategy that is good for all $t \geq T$, if the set is not approachable).

The definitions imply that every set approachable by one of the players is weakly approachable by that player, and every set excludable by one of the players is weakly excludable by that player. Vieille [1992] shows that every set is either weakly approachable by Player 1 or weakly excludable by Player 2. Compare this to Theorem 14.25 (page 588), with respect to the concepts of approachable and excludable sets, where a similar statement holds only for closed and convex sets.

In Section 14.7 (page 590), we used repeated games with vector payoffs to analyze a repeated game with incomplete information on one side. A similar analysis for general two-player zero-sum games with incomplete information for Player 2 and a finite set of states of nature is conducted by Aumann and Maschler [1995], who use repeated games with vector payoffs to show that the uniform value $v(p)$ of a repeated game with incomplete information for Player 2 is the smallest concave function that is greater than or equal to $u(p)$, where $u(p)$ is the value of the one-stage game in which the state of nature is chosen by the distribution p, and Player 1 does not make use of the information he has regarding the state of nature (equivalently, this is the game in which neither player knows the true state of nature). They also show that a strategy guaranteeing Player 2 the payoff $v(p) + \varepsilon$ is the strategy that approaches a proper set in an auxiliary repeated game with vector payoffs. They also construct a simple strategy for Player 1, guaranteeing him $v(p)$ (Exercise 14.29).

A geometric characterization of equilibria in two-player repeated games with incomplete information on one side that are not zero-sum appears in Hart [1986] (see also Aumann and Hart [1986]), and the existence of equilibria in these games is proved by Sorin [1983] for games with two states of nature, and by Simon, Spież, and Toruńczyk [1995], for games with an arbitrary number of states of nature. The existence of the value and equilibria in games with different information structures is studied in Kohlberg and Zamir [1974], Forges [1982], and Neyman and Sorin [1997, 1998], among others.

The first study of no-regret strategies was conducted by Hannan [1957]. The connection between no-regret strategies and the concept of approachable sets was first made by Hart and Mas-Colell [2000]. Several studies, including Foster and Vohra [1997] and Fudenberg and Levine [1999], define no-regret in a stronger form than the one presented here. Rustichini [1999], Lugosi, Mannor, and Stoltz [2007], and Lehrer and Solan [2007] studied no-regret strategies under which the decision maker does not know the true state of nature, but receives information that depends on the state of nature and the chosen action. No-regret strategies and their applications are covered in detail in Cesa-Bianchi and Lugosi [2006].

A characterization of approachable sets and excludable sets, where one of the players is restricted to strategies with finite memories, is given by Lehrer and Solan [2006, 2008] (such strategies are mentioned in Exercise 13.38, on page 563).

14.10 Remarks

A definition equivalent to that of a B-set, for convex sets, is given in Hart and Mas-Colell [2000].

Exercise 14.23 is based on Lehrer [2002], who calls the condition appearing there “the principle of approachability.” Exercise 14.28 is based on an example in Chapter I.3 in

Aumann and Maschler [1995], and Exercise 14.29 is based on Example IV.4.1 in that book. Exercise 14.30 is based on Zamir [1992]. Exercise 14.36 is based on Lehrer and Solan [2007].

14.11 Exercises

14.1 Let (X, d) be a metric space and let C be a subset of X. Prove that for each $x, y \in X$,

$$d(x, C) \leq d(x, y) + d(y, C). \tag{14.117}$$

14.2 Prove Markov's inequality: for every nonnegative random variable X, and every $c > 0$,

$$\mathbf{P}(X \geq c) \leq \frac{\mathbf{E}(X)}{c}. \tag{14.118}$$

14.3 Describe the following situation as a repeated game with vector payoffs, where Player 1 is M. Goriot, and Player 2 represents his daughters. At the start of every month, Anastasia and Delphine decide how to relate to their father, M. Goriot, that month: will they ignore his existence or pay him a visit now and again? M. Goriot, for his part, decides whether to give his daughters a generous, or a stingy, monthly allowance. If M. Goriot decides to be generous, he gives Anastasia 10 francs at the end of the month, and Delphine 12 francs, if his daughters have not visited him that month; and he gives Anastasia 18 francs at the end of the month, and Delphine 16 francs, if they have visited him. If M. Goriot decides to be stingy, he gives Anastasia 3 francs at the end of the month, and Delphine 2 francs, if his daughters have not visited him that month; and he gives Anastasia 5 francs at the end of the month, and Delphine 8 francs, if they have visited him.

14.4 Prove the following two claims:

(a) If strategy σ_k approaches a set C for player k, then it approaches the closure of C for that player.
(b) If strategy σ_k excludes a set C for player k, then it excludes the closure of C for that player.

14.5 Prove that the following two claims hold for any closed set $C \subseteq \mathbb{R}^m$ (recall that M is the maximal payoff of the game, in absolute value):

(a) C is approachable by a player if and only if the set $\{x \in C : \|x\| \leq M\}$ is approachable by the other player.
(b) C is excludable by a player if and only if the set $\{x \in C : \|x\| \leq M\}$ is excludable by the player.
(c) Show that if C is not closed, item (a) above does not necessarily hold.

14.6 Prove the following claims:

(a) If strategy σ_k approaches a set C for player k, then it approaches every superset of C for that player.
(b) If strategy σ_k excludes a set C for player k, then it excludes every subset of C for that player.

14.7 Prove Theorem 14.8 on page 575: a set cannot be both approachable by one player and excludable by the other player.

14.8 In this exercise we will prove that the set C_3 defined in Example 14.4, page 574 (see Figure 14.4) is approachable by Player 1, using the strategy described there.

(a) Prove that if $\overline{g}^{t-1}$ is above the diagonal $x_1 + x_2 = 1$, then $d(\overline{g}^t, \vec{0}) \leq d(\overline{g}^{t-1}, \vec{0}) - \frac{1}{2t}$.
(b) Prove that if $\overline{g}^{t-1}$ is under the diagonal $x_1 + x_2 = 1$, then $d(\overline{g}^t, C_2) \leq (1 - \frac{1}{2t}) d(\overline{g}^{t-1}, C_2)$.
(c) Using the fact that the series $\sum_{t=1}^{\infty} \frac{1}{t}$ diverges, and that the sequence $\{\frac{1}{t}\}$ converges to 0, deduce that $\lim_{t \to \infty} d(\overline{g}^t, C_3) = 0$.

14.9 For each of the following sentences, find an example in which the claim of the sentence obtains:

(a) The sets C_1 and C_2 are approachable by Player 1, but the set $C_1 \cap C_2$ is excludable by Player 2.
(b) The set C_1 is approachable by Player 1, the set C_2 is excludable by Player 2, and the set $C_1 \setminus C_2 := C_1 \cap (C_2)^c$ is excludable by Player 2.
(c) The sets C_1 and C_2 are excludable by Player 2, but the set $C_1 \cup C_2$ is approachable by Player 1.

14.10 Consider the following two-player game, with payoffs in $\mathbb{R}^2$:

Player 2

		L	R
Player 1	T	(−1, 1)	(1, −1)
	B	(2, 2)	(−1, −1)

(a) Draw the sets $R_1(p)$ and $R_2(q)$, for each p and q.
(b) Prove that the following two sets are approachable by Player 1:

$$C_1 = [(1, -1), (-1, 1)], \qquad (14.119)$$

$$C_2 = [(0, 0), (2, 2)] \cup [(0, 0), (1, -1)]. \qquad (14.120)$$

(c) Prove that the following three sets are not approachable by Player 1:

$$C_3 = [(-1, 1), (2, 2)], \tag{14.121}$$

$$C_4 = [(0, 0), (-1, 1)] \cup [(0, 0), (-1, -1)], \tag{14.122}$$

$$C_5 = \left\{\left(x, \tfrac{1}{2}\right) : -\infty < x < \infty\right\}. \tag{14.123}$$

14.11 Let C_1 be a convex set and let C_2 be a convex set containing C_1. Prove that if C_1 is a B-set for a certain player, then C_2 is also a B-set for that player.

14.12 In this exercise, we will show that Exercise 14.11 does not hold without the condition that C_1 and C_2 are convex sets. Consider the game presented in Exercise 14.10.

(a) Prove that the set $C_1 = [(0, 0), (2, 2)] \cup [(0, 0), (1, -1)]$ is a B-set for Player 1.
(b) Prove that for $\varepsilon > 0$ sufficiently small, the set C_ε, which is the union of C_1 with the triangle whose vertices are $(0, 0)$, $(\varepsilon, \varepsilon)$, and $(\varepsilon, -\varepsilon)$, is not a B-set for Player 1.

14.13 Answer the following questions for each one of the games below, whose payoffs are in $\mathbb{R}^2$.

(a) Draw the sets $R_1(p)$ and $R_2(q)$, for every p and q.
(b) Find four B-sets for Player 1.
(c) Find four B-sets for Player 2.

Player 2

Player 1

	L	R
T	(0, 1)	(0, 4)
B	(2, −1)	(1, 1)

Game A

Player 2

Player 1

	L	R
T	(0, 0)	(2, 2)
B	(3, 1)	(1, 1)

Game B

14.14 Answer the following questions for the game below, whose payoffs are in $\mathbb{R}^2$.

Player 2

Player 1

	L	R
T	(1, 0)	(0, 4)
B	(0, 2)	(3, 1)

(a) Draw the sets $R_1(p)$ and $R_2(q)$, for every p and q.
(b) Which of the following sets is a B-set for Player 1, which one is a B-set for Player 2, and which one is neither? Justify your answers.

- $C_1 = [(1, 0), (0, 4)]$.
- $C_2 = [(0, 4), (\frac{6}{11}, \frac{20}{11})] \cup [(0, 2), (\frac{6}{11}, \frac{20}{11})]$.

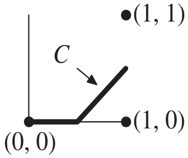
- $C_3 = [(3, 1), (\frac{6}{11}, \frac{20}{11})] \cup [(1, 1), (\frac{6}{11}, \frac{20}{11})]$.
- $C_4 = \{(\frac{6}{11}, \frac{20}{11})\}$.

14.15 Theorem 14.14 (on page 579) states that every B-set for a player is also an approachable set for that player, and Theorem 14.47 (page 606) states that every approachable set for a player contains a B-set for that player. In this exercise, we show that an approachable set for a player may not be a B-set for that player.

Consider the game appearing in Example 14.4 (page 574).

(a) Prove that the set C, which is the union of two intervals, $[(0, 0), (\frac{1}{2}, 0)]$ and $[(\frac{1}{2}, 0), (1, \frac{1}{2})]$ (see the accompanying figure), is not a B-set for Player 1.

(b) Find a B-set for Player 1 contained in C, and deduce that C is indeed an approachable set for Player 1. Prove that the set that you have found is a B-set for Player 1.



14.16 In this exercise we will prove von Neumann's Theorem (Theorem 5.11 on page 151) using results proved in this chapter. Let $\Gamma = (N, S_1, S_2, u)$ be a two-player zero-sum game. Denote by $\underline{v}_1 = \max_{\sigma_1 \in \Delta(S_1)} \min_{\sigma_2 \in \Delta(S_2)} U_1(\sigma_1, \sigma_2)$ (respectively $\overline{v}_1 = \min_{\sigma_2 \in \Delta(S_2)} \max_{\sigma_1 \in \Delta(S_1)} U_1(\sigma_1, \sigma_2)$) the maxmin (respectively minmax) value in mixed strategies. Consider the game to be a game with one-dimensional vector payoffs, and define the sets $C = [\overline{v}_1, \infty)$ and $D = (-\infty, \underline{v}_1]$.

(a) Prove (without using von Neumann's Theorem) that C is approachable by Player 1.

(b) Prove (without using von Neumann's Theorem) that D is approachable by Player 2.

(c) Using the fact that $\overline{v}_1 \geq \underline{v}_1$ (Exercise 4.34 on page 137), deduce that $\overline{v}_1 = \underline{v}_1$.

14.17 Let $(Y_i)_{i=1}^{\infty}$ be a sequence of random variables, $Y_i \in [0, 1]$ for every $i \in N$. Let $(X_i)_{i=1}^{\infty}$ be a sequence of independent random variables, with Bernoulli distribution with parameter p, where $p \in (0, 1)$. In other words, $\mathbf{P}(X_i = 1) = p$, and $\mathbf{P}(X_i = 0) = 1 - p$.

(a) Prove that[10]

$$\mathbf{P}\left(\lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i - \frac{\sum_{\{i\,:\,X_i=1\}} Y_i}{|\{i : X_i = 1\}|}\right) = 0\right) = 1. \qquad (14.124)$$

In words, for n sufficiently large, the average of the first n elements in the sequence $(Y_i)_{i=1}^{\infty}$ is close to the average of independently chosen elements: every element is chosen with probability p.

---

[10] Recall that for every finite set X we denote by $|X|$ the number of elements it contains.

(b) Prove that Equation (14.124) holds, even when Y_i depends on $X_1, X_2, \dots, X_{i-1}$ (but is independent of $(X_j)_{j \geq i}$) for each $i \in \{1, 2, \dots, n\}$.

Guidance: Consider the following game with vector payoffs in $\mathbb{R}^4$:

$$\begin{array}{cc|c|c|} & & \multicolumn{2}{c}{\text{Player 2}} \\ & & 1 & 0 \\ \cline{3-4} \text{Player 1} & 1 & (1,0,0,0) & (0,1,0,0) \\ \cline{3-4} & 0 & (0,0,1,0) & (0,0,0,1) \\ \cline{3-4} \end{array}$$

Interpret Y_n as the mixed action of Player 2 in stage n, and X_n as the pure action of Player 1 in stage n. Prove that the strategy of Player 1, under which he plays action 1 with probability p, and action 0 with probability $1 - p$ (independently of previous choices) in each stage, approaches the set

$$C := \{(yp, (1-y)p, y(1-p), (1-y)(1-p)) \colon 0 \leq y \leq 1\} \tag{14.125}$$
$$= \{y(p, 0, 1-p, 0) + (1-y)(0, p, 0, 1-p) \colon 0 \leq y \leq 1\}. \tag{14.126}$$

Prove that the fact that C is approachable by Player 1 using the above-described strategy implies Equation (14.124) (in item (a)). Deduce the claim in item (b) from this.

14.18 Prove that Blackwell's Theorem (Theorem 14.14 on page 579) obtains even when for every $i \in \mathcal{I}$ and $j \in \mathcal{J}$, the payoff is a random variable $X(i, j)$ taking values in $\mathbb{R}^m$ with expectation $u(i, j)$, and Player 1 observes, after every stage t, the value of $X(i^t, j^t)$ and the action j^t of Player 2.

14.19 State the analogous corollary to Corollary 14.26 on page 588 for Player 2. Justify your answer.

14.20 Prove that a compact and convex set $C \subseteq \mathbb{R}^m$ is approachable by a player if and only if $C \cap H_2(q) \neq \emptyset$ for every $q \in \Delta(\mathcal{J})$.

14.21 Prove Corollary 14.26 for closed and convex sets that are not compact: a compact set C is approachable by Player 1 if and only if

$$\operatorname{val}(G_\alpha) \geq \inf_{x \in C} \langle \alpha, x \rangle, \quad \forall \alpha \in \mathbb{R}^m. \tag{14.127}$$

14.22 Consider the following two-player game with payoffs in $\mathbb{R}^2$:

$$\begin{array}{cc|c|c|} & & \multicolumn{2}{c}{\text{Player 2}} \\ & & L & R \\ \cline{3-4} \text{Player 1} & T & (0,1) & (1,0) \\ \cline{3-4} & B & (1,1) & (0,0) \\ \cline{3-4} \end{array}$$

(a) Write down the payoff function U_π in the game G_π for every $\pi = (\pi_1, \pi_2) \in \mathbb{R}^2$ (for the definition of U_π see Equation (14.68) on page 588).

(b) Draw the graph of the function $\mathrm{val}(G_\pi)$. That is, on the two-dimensional plane, where the x axis is identified with π_1, and the y axis is identified with π_2, draw $\mathrm{val}(G_{\pi_1,\pi_2})$ at each point.

(c) For the following sets C, compute the value of $\min_{x \in C} \langle \pi, x \rangle$ as a function of π, and determine which of them are approachable by Player 1.

  (i) $C_1 = \{(\frac{1}{2}, \frac{1}{2})\}$.

  (ii) $C_2 = [(0, 1), (1, 0)]$.

  (iii) $C_3 = [(0, 1), (0, 0)]$.

  (iv) C_4 is the triangle whose vertices are $(1, 1)$, $(1, 0)$, and $(\frac{1}{2}, \frac{1}{2})$.

  (v) C_5 is the parallelogram whose vertices are $(\frac{1}{2}, 1)$, $(\frac{1}{4}, 1)$, $(\frac{1}{2}, 0)$, and $(\frac{3}{4}, 0)$.

14.23 **The principle of approachability** In this exercise, we will present the geometric principle behind Blackwell's Theorem. Let $C \subset \mathbb{R}^m$ be a compact and convex set, and let $(x^t)_{t \in \mathbb{N}}$ be a sequence of vectors in $\mathbb{R}^m$. Denote by $\overline{x}^t = \frac{1}{t}\sum_{j=1}^{t} x^j$ the average of the first t elements in the sequence $(x^t)_{t \in \mathbb{N}}$ and denote by $y(\overline{x}^t, C)$ the point in C that is closest to $\overline{x}^t$. Assume that for each t, the following inequality holds:

$$\langle \overline{x}^t - y(\overline{x}^t, C), x^{t+1} - y(\overline{x}^t, C) \rangle \leq 0, \quad \forall t \in \mathbb{N}. \tag{14.128}$$

(a) Prove that if $\overline{x}^t \notin C$, then Equation (14.128) holds if and only if the hyperplane tangent to C at $y(\overline{x}^t, C)$ separates x^{t+1} from $\overline{x}^t$.

(b) Prove that $\lim_{t \to \infty} d(\overline{x}^t, y(\overline{x}^t, C)) = 0$: the distance of the average $\overline{x}^t$ from C converges to 0.

(c) Explain why this exercise generalizes Blackwell's Theorem for compact and convex sets.

Guidance: See the last part of the proof of Lemma 14.15. For item (c) substitute $x^t = u(i^t, j^t)$.

14.24 Let C be a compact and convex set in $\mathbb{R}^m$. Let $(x^t)_{t=1}^{\infty}$ be a sequence of points in C satisfying

$$\langle \overline{x}^t, x^{t+1} \rangle \leq \langle \overline{x}^t, z \rangle, \quad \forall z \in C. \tag{14.129}$$

Prove that $\lim_{t \to \infty} d(\overline{x}^t, \mathrm{argmin}_{z \in S} \langle z, z \rangle) = 0$: the sequence of averages $(\overline{x}^t)_{t=1}^{\infty}$ converges to the point minimizing $\langle z, z \rangle$ in C.

Guidance: Show, using the convexity of C, that the set $\mathrm{argmin}_{z \in C} \langle z, z \rangle$ contains a single point y. This is the point in C closest to the origin. Apply the principle of approachability (Exercise 14.23) to the set $\mathrm{argmin}_{z \in C} \langle z, z \rangle$.

14.25 Let $M > 0$. For every $k \in \mathbb{N}$, let $u(k) : \mathcal{I} \times \mathcal{J} \to [-M, M]^m$ be a vector payoff function, and let $G(k)$ be the repeated game whose payoff function is $u(k)$. Let $\widehat{G}$ be the game in which the payoffs change from one stage to another: the payoffs in stage k are given by the function $u(k)$. Answer the following questions:

(a) Let $C \subseteq \mathbb{R}^m$ be a convex and closed set. Prove that if for every $k \in \mathbb{N}$ the set C is a B-set for Player 1 in $G(k)$, then C is approachable by Player 1 in the game $\widehat{G}$.
(b) Show by example that the result in the previous item does not necessarily hold if C is not convex.

14.26 Consider the game described in Section 14.7 (page 590).

(a) Prove that the value of the one-stage game is $\frac{1}{2}$.
(b) Prove the value of the two-stage game is $\frac{3}{8}$.
(c) Denote by v^T the value of the T-stage repeated game. Prove that $\lim_{T\to\infty} v^T = \frac{1}{4}$.

Guidance: For item (c), use the fact that the uniform value of the game (Definition 14.28 on page 592) is $\frac{1}{4}$.

14.27 This exercise presents a generalization of the repeated game with incomplete information for Player 2 in Section 14.7 (page 590) to the case where the state of nature s is chosen by a probability distribution: $\mathbf{P}(s = 1) = p$ and $\mathbf{P}(s = 2) = 1 - p$, where $p \in [0, 1]$. Denote by $v(p)$ the uniform value of the game. We will prove that $v(p) = p(1 - p)$ for every $p \in [0, 1]$.

(a) Prove that $v(0) = v(1) = 0$.
(b) Prove that Player 1, knowing s, can guarantee $p(1 - p)$ by not using his information, and playing the same mixed action in every stage. What mixed action will he choose?
(c) Define $C := \operatorname{conv}\{(0, 0), (\frac{1-p}{2}, 0), (0, \frac{p}{2}), (\frac{1-p}{2}, \frac{p}{2})\}$. Prove that if C is approachable by Player 2, then $v(p) \leq p(1 - p)$.
(d) Prove that C is approachable by Player 2.
(e) Deduce that $v(p) = p(1 - p)$ for every $p \in [0, 1]$.
(f) For every $\varepsilon > 0$, find a strategy for Player 2 that guarantees $p(1 - p) + \varepsilon$.

14.28 In this exercise, we present an example of a repeated game with incomplete information for Player 2, in which, under the optimal strategy, Player 1 reveals all the information in his possession. Consider the repeated game with incomplete information, as defined in Section 14.7 (page 590), with the following state games:

		Player 2	
		L	R
Player 1	T	−1	0
	B	0	0

The state game $s = 1$

		Player 2	
		L	R
Player 1	T	0	0
	B	0	−1

The state game $s = 2$

The state of nature is chosen using the distribution: $\mathbf{P}(s = 1) = p$ and $\mathbf{P}(s = 2) = 1 - p$, where $0 \leq p \leq 1$.

(a) Denote by $w(p)$ the value of the game in which Player 1 is restricted to playing the strategies in $\mathcal{B}_1^*$; in other words, he does not make use of his information regarding the state of nature. Prove that $w(p) = -p(1-p)$.

(b) For each player, find a strategy guaranteeing him 0 in the repeated game. Deduce that the uniform value of the game is $v(p) = 0$, for every $p \in [0, 1]$, and explain how Player 1 uses the information in his possession.

14.29 Repeat Exercise 14.28, using the following payoff functions:

		Player 2	
		L	R
Player 1	T	2	0
	B	0	1

The state game $s = 1$

		Player 2	
		L	R
Player 1	T	1	0
	B	0	2

The state game $s = 2$

(a) Prove that $w(p) = \frac{(2-p)(1+p)}{3}$.

(b) Write out the corresponding repeated game with vector payoffs G_V.

(c) Sketch, in the plane, the sets F and $R_2(q)$ for $q \in [0, 1]$.

(d) Prove that for all $p \in [0, 1]$, the set

$$C := \operatorname{conv}\left\{(0,0), \left(0, \frac{p^2+2}{3}\right), \left(\frac{p^2-2p+3}{3}, 0\right), \left(\frac{p^2-2p+3}{3}, \frac{p^2+2}{3}\right)\right\} \tag{14.130}$$

is a B-set for Player 2 in the game G_V.

(e) What is $v(p)$, the uniform value of the repeated game, for every $p \in [0, 1]$? Justify your answer.

Guidance: To prove that C is a B-set for Player 2 in the game G_V, divide the points in F into three sets: $A_1 := \{x \in F : x_1 \leq \frac{p^2-2p+3}{3}\}$, $A_2 := \{x \in F : x_2 \leq \frac{p^2+2}{3}\}$, and $A_3 := F \setminus (A_1 \cup A_2)$. The q corresponding to $x \in A_1$ in the definition of the B-set (of Player 2) is $q = \frac{p^2+2}{3}$; the q corresponding to $x \in A_2$ in the definition of the B-set (of Player 2) is $q = \frac{p^2-2p+3}{3}$. For $x \in A_3$, compute the supporting line of C at the point $(\frac{p^2-2p+3}{3}, \frac{p^2+2}{3})$ that is the perpendicular to the line connecting x to that point, and show that there is a q such that this supporting line separates x from $R_2(q)$.

14.30 In this exercise, we present an example of a repeated game with incomplete information for Player 2 in which, under the optimal strategy, Player 1 reveals only some of the information in his possession. Consider the repeated game with incomplete information for Player 2, where Player 1 has two actions $\mathcal{I} = \{T, B\}$, Player 2 has three actions $\mathcal{J} = \{L, M, R\}$, and there are two states of nature $S = \{1, 2\}$. The state of nature is chosen by the toss of a fair coin: $\mathbf{P}(s = 1) = \mathbf{P}(s = 2) = \frac{1}{2}$, and the state games are given by the following tables:

		Player 2		
		L	M	R
Player 1	T	4	0	2
	B	4	0	–2

The state game $s = 1$

		Player 2		
		L	M	R
Player 1	T	0	4	–2
	B	0	4	2

The state game $s = 2$

(a) Prove that when Player 1 does not use the information in his possession, he can guarantee at most 0; that is, the value of the game in which Player 1's set of strategies is $\mathcal{B}_1^*$, and Player 2's set of strategies is $\mathcal{B}_2$, is 0.

(b) Prove that when Player 2 also knows the state of nature, the value of the game is 0.

(c) Returning to the game in which only Player 1 knows the state of nature, define the strategy $\widehat{\sigma}_1$ of Player 1 as follows:

- In the first stage, if $s = 1$, play the mixed action $[\frac{3}{4}(T), \frac{1}{4}(B)]$. If $s = 2$, play the mixed action $[\frac{1}{4}(T), \frac{3}{4}(B)]$.
- In each stage $t > 1$, repeat the action that was played in the first stage; i.e., play either T in every stage, or B in every stage, according to the result of the lottery in the first stage.

Compute the conditional probability $\mathbf{P}(s = 1 \mid i^1 = T)$ that the state of the world is $s = 1$, given that Player 1 played T in the first stage. Compute also $\mathbf{P}(s = 1 \mid i^1 = B)$.

(d) Show that for every strategy of Player 2, the conditional expectation of the average payoff $\overline{g}^t$, given the action of Player 1 in the first stage, is at least 1. Deduce that Player 1 can guarantee at least 1 in the infinitely repeated game.

(e) Prove that if the set $C = [0, 1]^2$ is approachable by Player 2, then the uniform value of the game is at most 1.

(f) Draw, in the plane, the set of possible payoffs F and the sets $R_2(0, 0, 1)$, $R_2(0, 1, 0)$, $R_2(1, 0, 0)$, $R_2(\frac{1}{2}, \frac{1}{2}, 0)$, and $R_2(\frac{1}{2}, 0, \frac{1}{2})$.

(g) Prove that C is approachable by Player 2. Note that when checking the condition in the definition of a B-set, one needs to distinguish between points on the diagonal $x_1 = x_2$ and points below the diagonal.

(h) In which stages of the game does Player 1 reveal information on the state of nature? Does he ever entirely reveal the state of nature?

14.31 In this exercise, we generalize the results presented in Section 14.7 (page 590), and Exercises 14.28, 14.29, and 14.30, to two-player zero-sum infinitely repeated games with incomplete information on one side. Consider the following game:

- There are K state games (states of nature). In all state games, the set of actions of Player 1 is $\mathcal{I} = \{1, 2, \dots, I\}$, and the set of actions of Player 2 is $\mathcal{J} = \{1, 2, \dots, J\}$. The payoff function in state game k is u_k; $u_k(i, j)$ is the payoff that Player 2 pays Player 1 when the state of nature is the matrix u_k, and the pair of actions chosen is (i, j). Denote $S = \{1, 2, \dots, K\}$.

- The game begins with a move of chance that selects one of the payoff matrices according to the probability distribution $p = (p_k)_{k=1}^K$, which is common knowledge among the players.
- Player 1 knows which state game has been chosen, but Player 2 does not have this information.
- In each stage, the two players choose their actions simultaneously: Player 1 chooses an action in $\mathcal{I}$, and Player 2 chooses an action in $\mathcal{J}$.
- In each stage t, when coming to choose an action, each player knows the actions chosen by both players in the previous stages. The players are not informed of their payoffs after each stage (although Player 1, knowing the state game and the actions chosen, can calculate the payoffs, while Player 2 cannot do so).
- This description of the game is common knowledge among the players.

For every distribution $p \in \Delta(S)$, denote the game described above by $G(p)$, denote by $v(p)$ the value of $G(p)$, if the game has a value, and denote by $D(p)$ the one-stage game in which the payoff matrix is chosen according to the distribution p, and neither Player 1 nor Player 2 knows which matrix has been chosen. The game $D(p)$ is a one-stage game, and therefore its value exists. Denote this value by $w(p)$.

Let cav w be the smallest concave function that is pointwise greater than or equal to w. In other words:

$$(\text{cav } w)(q) \geq w(q), \quad \forall q \in \Delta(S). \tag{14.131}$$

In this exercise, we will prove that the value $v(p)$ exists, and equals $(\text{cav } w)(p)$, and we will construct the players' optimal strategies.

We first prove that Player 1 can guarantee this value.

(a) Show that for every $p \in \Delta(S)$, Player 1 can guarantee $w(p)$ in $G(p)$ (find a strategy of Player 1 that guarantees this value).

(b) Let $p^1, p^2, \dots, p^L \in \Delta(S)$, and let $\alpha = (\alpha_l)_{l=1}^L$ be a distribution over $\{1, 2, \dots, L\}$ satisfying $p = \sum_{l=1}^L \alpha_l p^l$. Consider the following strategy σ_1^* of Player 1: if u_{k_0} is the payoff function that has been selected in the chance move at the beginning of the game, randomly choose $l_0 \in \{1, 2, \dots, L\}$ according to the following distribution $\lambda^{k_0} = (\lambda_1^{k_0}, \lambda_2^{k_0}, \dots, \lambda_L^{k_0})$:

$$\lambda_l^{k_0} = \frac{\alpha_l p_{k_0}^l}{\sum_{l'=1}^L \alpha_{l'} p_{k_0}^{l'}} = \frac{\alpha_l p_{k_0}^l}{p_{k_0}}. \tag{14.132}$$

In each stage of the game, play the optimal strategy in the game $D(p^{l_0})$.

We will prove that this strategy guarantees Player 1 the payoff $\sum_{l=1}^L \alpha_l w(p^l)$ in $G(p)$.

To accomplish this, consider the game $G'(p)$, a variation of $G(p)$, where Player 2 is informed of the index l_0 that Player 1 chose. Show that after Player 2 knows l_0, the conditional probability that the chosen payoff matrix is u_k, given

l_0, equals $p_k^{l_0}$. Deduce from this that in $G'(p)$, the strategy σ_1^* guarantees Player 1 the[11] payoff $\sum_{l=1}^L \alpha_l w(p^l)$.

(c) Using the fact that in the game $G(p)$, Player 2 has fewer strategies than he has in $G'(p)$ (because he has more information in $G'(p)$), deduce, with the help of Exercise 4.27 (page 135), that in $G(p)$, the strategy σ_1^* guarantees Player 1 the payoff $\sum_{l=1}^L \alpha_l w(p^l)$.

(d) Deduce that Player 1 can guarantee $(\text{cav } w)(p)$ in $G(p)$.

We now prove that Player 2 can also guarantee $(\text{cav } w)(p)$ in $G(p)$. Define an auxiliary game with vector payoffs G_V, where the sets of actions of the players are $\mathcal{I}$ and $\mathcal{J}$ respectively, and the matrix of vector payoffs is

$$u(i,j) = (u_1(i,j), u_2(i,j), \ldots, u_K(i,j)) \in \mathbb{R}^K. \tag{14.133}$$

Define

$$Z := \{z \in \mathbb{R}^K : \langle q, z\rangle \geq w(q), \quad \forall q \in \Delta(S)\}. \tag{14.134}$$

The significance of Z stems from the following claim:

(e) Prove that $(\text{cav } w)(p) = \min_{z\in Z}\langle p, z\rangle$.

For every $z \in Z$, define

$$M_z := z - \mathbb{R}_+^K = \{x \in \mathbb{R}^K : x_k \leq z_k \quad \forall k \in S\}. \tag{14.135}$$

As we will see, the set M_z is approachable by Player 2 in G_V, and this is the set corresponding to the rectangles C defined in Equation (14.85) (page 596), and in Exercises 14.29 and 14.30.

(f) Prove that the set M_z is approachable by Player 2 in G_V if and only if for every $\pi \in \mathbb{R}^K$,

$$\max_{y\in\Delta(\mathcal{J})} \min_{x\in\Delta(I)} U_\pi(x,y) \geq \min_{z'\in M_z}\langle \pi, z'\rangle. \tag{14.136}$$

Guidance: use Corollary 14.26 (page 588), reversing the roles of the players.

(g) Prove that Equation (14.136) obtains for $\pi = \vec{0}$.

(h) Let $\pi \neq \vec{0}$ satisfy the property that there exists $k \in S$ such that $\pi_k > 0$. Prove that the right-hand side of Equation (14.136) equals $-\infty$, while the left-hand side is finite, and therefore Equation (14.136) obtains.

(i) We now show that $z \in Z$ if and only if Equation (14.136) holds for every $\pi \neq \vec{0}$ satisfying $\pi_k \leq 0$ for all $k \in S$. Let $\pi \neq \vec{0}$ satisfy $\pi_k \leq 0$ for all $k \in S$. Define $\widehat{q}_k = -\frac{\pi_k}{\sum_{k'=1}^K \pi_{k'}}$. Show that $\widehat{q} = (\widehat{q}_k)_{k=1}^K \in \Delta(S)$.

---

[11] The lottery that chooses l_0 has the property that the conditional distribution over S, after the lottery, will be one of the probability distributions $p^1, p^2, \ldots, p^L$ with probabilities $\alpha_1, \alpha_2, \ldots, \alpha_L$, respectively. It follows that if for any l_0 chosen, Player 1 can guarantee $w(p^{l_0})$, then he can guarantee $\sum_{l=1}^L \alpha_l w(p^l)$ in $G(p)$. This random choice is the only stage in which Player 1 uses the knowledge he has regarding the payoff matrix that has been chosen. After l_0 has been chosen, Player 1 plays, in every stage, an optimal strategy in $G(p^{l_0})$, independently of the true payoff matrix u_{k_0}, and thus he does not reveal any additional information about the chosen payoff matrix in those stages.

(j) Prove that Equation (14.136) obtains for π if and only if the following equation obtains:

$$\min_{y\in\Delta(\mathcal{J})}\max_{x\in\Delta(I)} U_{\widehat{q}}(x,y) \leq \max_{z'\in M_z}\langle \widehat{q}, z'\rangle. \tag{14.137}$$

(k) Prove that the right-hand side of Equation (14.137) equals $\langle q, z\rangle$.

(l) Prove that the left-hand side of Equation (14.137) equals $w(\widehat{q})$.

(m) Deduce that $z \in Z$ if and only if M_z is approachable by Player 2 in G_V.

(n) Let $z_0 \in Z$ satisfy $\langle p, z_0\rangle = \min_{z\in Z}\langle p, z\rangle$. Deduce that M_{z_0} is approachable by Player 2 in G_V.

(o) Prove that the strategy of Player 2 that approaches M_{z_0} in G_V guarantees him $(\text{cav } w)(p)$ in $G(p)$.

14.32 We proved Theorem 14.44 (page 603) for the case in which $|E| = |A|$, and every expert recommends one action in every stage. Prove the theorem for any set E of experts.
Guidance: Use Exercise 14.25.

14.33 In this exercise we consider a decision problem *without* experts. A strategy σ of a decision maker in a decision problem is a *no-regret strategy* if for every action $\widehat{a} \in A$, and every sequence $(s^1, s^2, \ldots)$ of states of nature,

$$\mathbf{P}_\sigma\left(\liminf_{T\to\infty}\left(\frac{1}{T}\sum_{t=1}^{T} u(a^t, s^t) - \frac{1}{T}\sum_{t=1}^{T} u(\widehat{a}, s^t)\right) \geq 0\right) = 1. \tag{14.138}$$

Prove that there exists a no-regret strategy in every decision problem.
Guidance: Show that this problem may be described as a decision problem with experts, and use Theorem 14.44 (page 603) to prove the claim.

14.34 In this exercise, we present Blackwell's proof of Theorem 14.44 (page 603), which states that there exists a no-regret strategy when there are $|A|$ experts, and each of them is identified with one action (as was assumed in Section 14.8.2, on page 603).

Consider a decision problem with experts, where $E = A$, and for every action $a \in A$ there exists an expert $e \in E$ such that $a^t_e = a$ for every $t \in \mathbb{N}$. Define a game with vector payoffs, where the set of actions of Player 1 is A, the set of actions of Player 2 is S, and the payoffs are $(|S|+1)$-dimensional vectors $(w(s,a))^{s\in S}_{a\in A}$, where $w(s,a) = ((w_{s'}(s,a))_{s'\in S}, U(s,a))$, defined as follows:

$$\begin{aligned} w_{s'}(s,a) &= 1 \qquad s' = s,\\ w_{s'}(s,a) &= 0 \qquad s' \neq s. \end{aligned} \tag{14.139}$$

Define a set $C \subseteq \mathbb{R}^{|S|+1}$ as follows:

$$C := \{(q,x) \in \Delta(S)\times\mathbb{R} : x \geq \max_{p\in\Delta(A)} U(p,q)\}. \tag{14.140}$$

(a) Prove that C is a convex set.

(b) Prove that C is not excludable by Player 2.

(c) Deduce that C is approachable by Player 1, and prove that every strategy of Player 1 that approaches C is a no-regret strategy: if Player 1 plays this strategy, then with probability 1

$$\liminf_{t\to\infty} u(\overline{x}^t, \overline{y}^t) - \max_{a\in A} U(a, \overline{y}^t) \geq 0. \tag{14.141}$$

14.35 In this exercise, we present an alternative proof to the proof given in Exercise 14.34. Repeat Exercise 14.34 for the game where the set of actions of Player 1 is A, the set of actions of Player 2 is S, and the payoffs are $(|S| \times |A|)$-dimensional vectors $(w(s,a))_{a\in A}^{s\in S}$, with $w(s,a) = (w_{s',a'}(s,a))_{a'\in A}^{s'\in S}$ defined as follows:

$$\begin{aligned} w_{s',a'}(s,a) &= 1 \qquad (s',a') = (s,a), \\ w_{s',a'}(s,a) &= 0 \qquad (s',a') \neq (s,a), \end{aligned} \tag{14.142}$$

and for the following set C

$$C := \left\{ q \in \Delta(S\times A) \colon U(\mathbf{1}_a, q_{|a}) \geq \max_{p\in\Delta(A)} U(p, q_{|a}), \quad \forall a \in A \right\}, \tag{14.143}$$

where U is the multilinear extension of u, and for every action $a \in A$, $q_{|a}$ is the conditional distribution over S given a, with this conditional distribution defined by

$$q_{|a}(s) := \begin{cases} \frac{q(s,a)}{\sum_{s'\in S} q(s',a)} & \text{if } \sum_{s'\in S} q(s''a) > 0, \\ 0 & \text{if } \sum_{s'\in S} q(s',a) = 0. \end{cases} \tag{14.144}$$

14.36 Let S be a finite set of states of nature, and let A be a finite set of actions available to a decision maker. Let $F : \Delta(S) \to 2^A$ be a set-valued function associating each $y \in \Delta(S)$ with a set $F(y) \subseteq A$ and satisfying the property that for every $a \in A$, the set $F^{-1}(a) := \{y \in \Delta(S) \colon a \in F(y)\}$ is closed and convex. When the state of nature is chosen according to y, the best actions, from the perspective of the decision maker, are those in $F(y)$. Define a game with $(|A| \times |S|)$-dimensional vector payoffs as follows: the payoff $u(a,s)$ that is attained when the state of nature is s, and the action chosen is a, is the unit vector whose value is 0 at every coordinate except the coordinate (a,s), where its value is 1. Denote this vector by $\mathbf{1}_{as} \in \mathbb{R}^{|A|\times|S|}$. Define a set $C \subset \mathbb{R}^{|S|\times|A|}$ as follows:

$$C := \operatorname{conv}\left\{ \sum_{s\in S} y_s \mathbf{1}_{as}, a \in A, y \in F^{-1}(a) \right\}. \tag{14.145}$$

(a) Prove that the set C is a convex set.
(b) Prove that C is approachable by Player 1.
(c) Let n_a^t be the number of stages, up to t, in which the action a is chosen, and let y_a^t be the empirical distribution of the states of nature in all stages up to stage t in which the action a is chosen:

$$n_a^t := |\{j \leq t \colon a^j = a\}|, \quad y_a^t(s) := \frac{1}{n_a^t} |\{j \leq t \colon s^j = s, a^j = a\}|.$$

Prove that if the decision maker plays a strategy that approaches C, and if $\liminf_{t\to\infty} \frac{n_a^t}{t} > 0$, then $\lim_{t\to\infty} d(y_a^t, F^{-1}(a)) = 0$. In other words, if the density of the stages where the action a is chosen is positive, then the empirical distribution of the states of nature in the stages in which the action a is chosen approaches $F^{-1}(a)$, which is the set of distributions for which a is the best response in the decision maker's opinion.

14.37 Morris notices that in the proof of Theorem 14.45 (page 603), in order to prove that $\mathrm{val}(\widehat{G}_\pi) \geq \inf_{x\in C}\langle \pi, x\rangle$ for every $\pi \in \mathbb{R}^A$, the only fact that is used is that the vectors $\vec{0}$ and $(x^k)_{k\in\mathbb{N}}$ are in the nonnegative quadrant of $\mathbb{R}$. Morris therefore claims that every set containing these vectors is approachable in the auxiliary game. Is Morris correct? If not, what is his error?

14.38 Prove that a set cannot be weakly approachable by Player 1 and also weakly excludable by Player 2.