13

Repeated games

Chapter summary

In this chapter we present the model of repeated games. A repeated game consists of a base game, which is a game in strategic form, that is repeated either finitely or infinitely many times. We present three variants of this model:

- The finitely repeated game, in which each player attempts to maximize his average payoff.
- The infinitely repeated game, in which each player attempts to maximize his long-run average payoff.
- The infinitely repeated game, in which each player attempts to maximize his discounted payoff.

For each of these models we prove a *Folk Theorem*, which states that under some technical conditions the set of equilibrium payoffs is (or approximates) the set of feasible and individually rational payoffs of the base game.

We then extend the Folk Theorems to uniform equilibria for discounted infinitely repeated games and to uniform ε-equilibria for finitely repeated games. The former is a strategy vector that is an equilibrium in the discounted game, for every discount factor sufficiently close to 1, and the latter is a strategy vector that is an ε-equilibrium in all sufficiently long finite games.

In the previous chapters, we dealt with one-stage games, which model situations where the interaction between the players takes place only once, and once completed, it has no effect on future interactions between the players. In many cases, interaction between players does not end after only one encounter; players often meet each other many times, either playing the same game over and over again, or playing different games. There are many examples of situations that can be modeled as multistage interactions: a printing office buys paper from a paper manufacturer every quarter; a tennis player buys a pair of tennis shoes from a shop in his town every time his old ones wear out; baseball teams play each other several times every season. When players repeatedly encounter each other in strategic situations, behavioral phenomena emerge that are not present in one-stage games.

- The very fact that the players encounter each other repeatedly gives them an opportunity to cooperate, by conditioning their actions in every stage on what happened in previous

stages. A player can threaten his opponent with the threat "if you do not cooperate now, in the future I will take actions that harm you," and he can carry out this threat, thus "punishing" his opponent. For example, the manager of a printing office can inform a paper manufacturer that if the price of the paper he purchases is not reduced by 10% in the future, he will no longer buy paper from that manufacturer.
- Repeated games enable players to develop reputations. A sporting goods shop can develop a reputation as a quality shop, or a discount store.

In this chapter, we present the model of repeated games. This is a simple model of games in which players play the same base game time and again. In particular, the set of players, the actions available to the players, and their payoff functions do not change over time, and are independent of past actions. This assumption is, of course, highly restrictive, and it is often unrealistic: in the example above, new paper manufacturers enter the market, existing manufacturers leave the market, there are periodic changes in the price of paper, and the quantity of paper that printers need changes over time. This simple model, however, enables us to understand some of the phenomena observed in multistage interactions. The more general model, where the actions of the players and their payoff functions may change from one stage to another, is called the model of "stochastic games." The reader interested in learning more about stochastic games is directed to Filar and Vrieze [1997] and Neyman and Sorin [2003].

13.1 The model

A repeated game is constructed out of the base game Γ that defines it, i.e., the game that the players play at each stage. We will assume that the base game is given in strategic form $\Gamma = (N, (S_i)_{i\in N}, (u_i)_{i\in N})$, where $N = \{1, 2, \dots, n\}$ is the set of players, S_i is the set of actions[1] available to player i, and $u_i : S \to \mathbb{R}$ is the payoff function of player i in the base game, where $S = S_1 \times S_2 \times \cdots \times S_n$ is the set of action vectors.

In repeated games, the players encounter each other again and again, playing the same strategic-form game Γ each time. The complete description of a repeated game needs to include the number of stages that the game is played. In addition, since the players receive a payoff at each stage, we need to specify how the players value the sequence of payoffs that they receive, i.e., how each player compares each payoff sequence to another payoff sequence. We will consider three cases:

- The game lasts a finite number of stages T, and every player wants to maximize his average payoff.
- The game lasts an infinite number of stages, and every player wants to maximize the upper limit of his average payoffs.

---

[1] In this chapter we will call the elements of S_i "actions," and reserve the term "strategy" for strategies in the repeated game.

- The game lasts an infinite number of stages, and each player wants to maximize the time-discounted sum of his payoffs.

Denote by

$$M := \max_{i \in N} \max_{s \in S} |u_i(s)| \tag{13.1}$$

the maximal absolute value of the payoffs received by the players in one stage. Recall that the set of distributions over a set S_i is $\Sigma_i = \Delta(S_i)$, the product set of these sets is $\Sigma = \Sigma_1 \times \Sigma_2 \times \cdots \times \Sigma_n$, and $U_i : \Sigma \to \mathbb{R}$ is the multilinear extension of the payoff functions u_i (defined over S; see page 147).

By definition, a strategy instructs a player how to play throughout the game. The definition of a strategy in finite repeated games, and infinitely repeated games, will be presented when these games are defined.

13.2 Examples

The following example will be referenced often, for illustrating definitions, and explaining claims in this chapter.

Example 13.1 **Repeated Prisoner's Dilemma** Recall that the Prisoner's Dilemma is a one-stage two-player game, depicted in Figure 13.1.

		Player II	
		D	C
Player I	D	1, 1	4, 0
	C	0, 4	3, 3

Figure 13.1 The one-stage Prisoner's Dilemma

For both players, action D strictly dominates action C, so the only equilibrium of the base game is (D, D).

Consider the case in which the players play the Prisoner's Dilemma twice, and the second time the game is played, they both know which actions were chosen the previous time they played the game. When this situation is depicted as an extensive-form game (see Figure 13.2), the game tree has information sets representing the fact that at each stage the players choose their actions simultaneously. In Figure 13.2, the total payoff of each player in the two stages are indicated by the leaves of the game tree, where the upper number is the total payoff of Player I, and the lower number is the total payoff of Player II. In this figure, and several other figures in this chapter, the depicted tree "grows" from top to bottom, rather than left to right, for the sake of saving space on the page.

What are the equilibria of this game? A direct inspection reveals that the strategy vector in which the players repeat the one-stage equilibrium (D, D) at both stages is an equilibrium of the two-stage

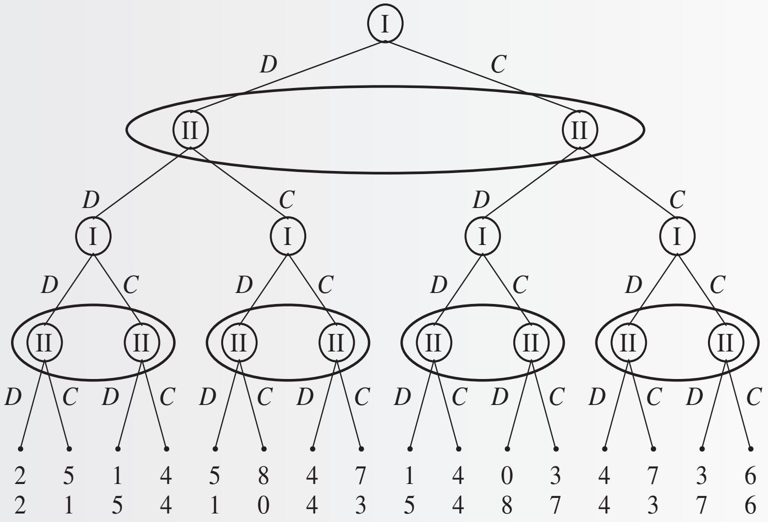

Figure 13.2 The two-stage Prisoner's Dilemma, represented as an extensive-form game

game. This is a special case of a general claim that states that every strategy vector where in every stage the players play an equilibrium of the base game is an equilibrium of the T-stage game (Theorem 13.6).

We argue now that at every equilibrium of the two-stage repeated game, the players play (D, D) in both stages. To see this, suppose instead that there exists an equilibrium at which, with positive probability, the players do not play (D, D) at some stage. Let $t \in \{1, 2\}$ be the last stage in which there is positive probability that the players will not play (D, D), and suppose that in this event, Player I does not play D in stage t. This means that if the game continues after stage t the players will play (D, D). We will show that this strategy cannot be an equilibrium strategy.

Case 1: $t = 1$. Consider the strategy of Player I at which he plays D in both stages. We will show that this strategy grants him a higher payoff. Since D strictly dominates C, Player I's payoff rises if he switches from C to D in the first stage. And since, by assumption, after stage t the players play (D, D) (since stage t is the last stage in which they may not play (D, D)), Player I's payoff in the second stage was supposed to be 1. By playing D in the second stage, Player I's payoff is either 1 or 4 (depending on whether Player II plays D or C);[2] in either case, Player I cannot lose in the second stage. The sum total of Player I's payoffs therefore rises.

Case 2: $t = 2$. Consider the strategy of Player I at which he plays in the first stage what the original strategy tells him to play, and in the second stage he plays D. Player I's payoff in the first stage does not change, but because D strictly dominates C, his payoff in the second stage does increase. The sum total of Player I's payoffs therefore increases.

---

[2] Even if $t = 1$ is the last stage in which one of the players plays C with positive probability, it is still possible that if both players play D in the first stage, then Player II will play C in the second stage with positive probability. To see this, consider the following strategy vector. In the first stage, both players play C. In the second stage, Player I plays D, and Player II plays D if Player I played C in the first stage, and he plays C if Player I played D in the first stage. In this case, if neither player deviates, the players play (C, C) in the first stage, and (D, D) in the second stage; but if Player I plays D in the first stage, then Player II plays C in the second stage.

Note that despite the fact that at every equilibrium of the two-stage repeated game the players play (D, D) in every stage, it is possible that at equilibrium, the strategy C is used off the equilibrium path; that is, if a player does deviate from the equilibrium strategy, the other player may play C with positive probability. For example, consider the following strategy σ_1:

- Play D in the first stage.
- In the second stage, play as follows: if in the first stage the other player played D, play D in the second stage; otherwise play $[\frac{1}{8}(C), \frac{7}{8}(D)]$ in the second stage.

Direct inspection shows that the strategy vector (σ_1, σ_1), in which both players play strategy σ_1, is an equilibrium of the two-stage repeated game.

By the same rationale used here to show that in the two-stage repeated Prisoner's Dilemma at equilibrium the players play (D, D) in both stages, it can be shown that in the T-stage repeated Prisoner's Dilemma, at equilibrium, the players play (D, D) in every stage (Exercise 13.6). ◀

As we saw, in the finitely repeated Prisoner's Dilemma, at every equilibrium the players play (D, D) in every stage. Does this extend to every repeated game? That is, does every equilibrium strategy of a repeated game call on the players to play a one-stage equilibrium in every stage? The following example shows that the answer is negative: in general, the set of equilibria of repeated games is a much richer set.

Example 13.2 **Repeated Prisoner's Dilemma, with the possibility of punishment** Consider the two-player game given in Figure 13.3, where each player has three possible actions.

	D	C	P
D	1, 1	4, 0	−1, 0
C	0, 4	3, 3	−1, 0
P	0, −1	0, −1	−2, −2

Figure 13.3 The repeated Prisoner's Dilemma, with the possibility of punishment

This game is similar to the Prisoner's Dilemma in Example 13.1, with the addition of a third action P to each player, yielding low payoffs for both players. Note that action P (which stands for Punishment) is strictly dominated by action D, and therefore by Theorem 4.35 (page 109) we can eliminate it without changing the set of equilibria of the base game. After eliminating P for both players, we are left with the one-stage Prisoner's Dilemma, whose only equilibrium is (D, D). It follows that the only equilibrium of the base game in Figure 13.3 is (D, D).

As previously stated, when the players play an equilibrium of the base game in every stage, the resulting strategy vector is an equilibrium of the repeated game. It follows that in the two-stage repeated game in this example, playing (D, D) in both stages is an equilibrium. In contrast with the standard repeated Prisoner's Dilemma, there are additional equilibria in this repeated game. The strategy vector at which both players play the following strategy is an equilibrium:

- Play C in the first stage.
- If your opponent played C in the first stage, play D in the second stage. Otherwise, play P in the second stage.

If both players play this strategy, they will both play C in the first stage, and D in the second stage, and each player's total payoff will be 4 (in contrast to the total payoff 2 that they receive under the equilibrium of playing (D, D) in both stages). Since action D weakly dominates both of the other actions, no player can gain by deviating from D in the second stage alone. A player who deviates in the first stage from C to D gets a payoff of 4 in the first stage, but he will then get at most -1 in the second stage (because his opponent will play P in the second stage), and so in sum total he loses: his total payoff when he deviates is 3, which is less than his total payoff of 4 at the equilibrium. By deviating to P in the first stage, the deviator also loses.

This example illustrates that in a repeated game, the players can threaten each other, by adopting strategies that call on them to punish a player in later stages, if at some stage that player deviates from a particular action. The greater the number of stages in the repeated game, the greater opportunity players have to punish each other. In general, this increases the number of equilibria.

The last equilibrium in this example is not a subgame perfect equilibrium (see Section 7.1 on page 252), since the use of the action P is not part of an equilibrium in the subgame starting in the second stage. We will see later in this chapter that repeated games may have additional equilibria that are subgame perfect.

Note that there is a proliferation of pure strategies in repeated games, compared to one-stage games. For example, in the one-stage game in Figure 13.3, every player has three pure strategies, D, C, and P. In the two-stage game, every player has $3 \times 3^9 = 3^{10} = 59{,}049$ pure strategies: there are three actions available to the player in the first stage, and in the second stage his strategy is given by a function from the pair of actions played in the first stage, i.e., from $\{D, C, P\}^2$ to $\{D, C, P\}$. In the three-stage repeated game, every player has $3 \times 3^9 \times (3^{3^4}) = 3^{91}$ pure strategies: the number of possible strategies in the first two stages is as calculated above, and in the third stage the player's strategy is given by a function from $\{D, C, P\}^4$ to $\{D, C, P\}$: for every pair of actions that were played in the first two stages, the player needs to decide what to play in the third stage. ◀

In general, the size of each player's space of strategies grows super-exponentially with the number of stages in the repeated game (Exercise 13.1). This growth has two consequences. A positive consequence is that it leads to complex and interesting equilibria. In Example 13.2, we found an equilibrium that grants a higher average payoff to the two players than their payoff when they repeat the only equilibrium of the one-stage game. A negative consequence is that, due to the complications inherent in the proliferation of strategies, it becomes practically impossible to find all the equilibria of repeated games with many stages. For this reason, we will not attempt to compute all equilibria of repeated games. We will instead look for asymptotic results, as the number of repetitions grows; we will seek approximations to the set of equilibrium payoffs, without trying to find all possible equilibrium payoffs; and we will be interested in special equilibria that can easily be described.

13.3 The *T*-stage repeated game

In this section we will study the equilibria of a T-stage repeated game Γ_T that is based on a strategic-form game Γ. Our goal is to characterize the limit set of equilibrium payoffs as T goes to infinity. We will also construct, for each vector x in the limit set of equilibrium payoffs, and for each sufficiently large natural number T, an equilibrium in the T-stage repeated game that yields a payoff close to x.

13.3.1 Histories and strategies

Since players encounter each other repeatedly in repeated games, they gather information as the game progresses. The information available to every player at stage $t+1$ is the actions played by all the players in the first t stages of the game. We will therefore define, for every $t \geq 0$, the *set of t-stage histories* as

$$H(t) := S^t = \underbrace{S \times S \times \cdots \times S}_{t \text{ times}}. \tag{13.2}$$

For $t = 0$, we identify $H(0) := \{\emptyset\}$, where $\emptyset$ is the history at the start of the game, which contains no actions. A history in $H(t)$ will sometimes be denoted by h^t, and sometimes by $(s^1, s^2, \dots, s^t)$, where $s^j = (s_i^j)_{i \in N}$ is the vector of actions played in stage j.

A behavior strategy for player i is an action plan that instructs the player which mixed action to play after every possible history.

Definition 13.3 *A* behavior strategy *for player i in a T-stage game is a function associating a mixed action with each history of length less than T*

$$\tau_i : \bigcup_{t=0}^{T-1} H(t) \to \Sigma_i. \tag{13.3}$$

The set of behavior strategies of player i in a T-stage game is denoted by $\mathcal{B}_i^T$.

Equivalently, we can define a behavior strategy of player i as a sequence $\tau_i = (\tau_i^t)_{t=0}^{T-1}$ of functions, where $\tau_i^{t+1} : H(t) \to \Sigma_i$ instructs the player what to play in stage t, for each $t \in \{0, 1, \dots, T-1\}$.

Remark 13.4 *When a T-stage repeated game is depicted as an extensive-form game, a pure strategy is a function* $\tau_i : \bigcup_{t=0}^{T-1} H(t) \to S_i$. *A mixed strategy is a distribution over pure strategies (Definition 5.3 on page 147). We have assumed that every player knows which actions were played at all previous stages; i.e., every player has perfect recall (see Definition 6.13 on page 109). By Kuhn's Theorem (Theorem 6.16 on page 235) it follows that every mixed strategy is equivalent to a behavior strategy, and we can therefore consider only behavior strategies, which are more convenient to use in this chapter.* ◆

Example 13.1 (*Continued*) Consider the two-stage Prisoner's Dilemma. Two (behavior) strategies are written in Figure 13.4, one for each player. The notation $\tau_{\mathrm{I}}(DC) = \left[\frac{2}{3}(D), \frac{1}{3}(C)\right]$ means that after history DC (which occurs if in the first stage Player I plays D, and Player II plays C), Player I plays the mixed action $\left[\frac{2}{3}(D), \frac{1}{3}(C)\right]$ in the second stage.

$$\begin{array}{ll}
\tau_{\mathrm{I}}(\phi) = \left[\frac{1}{2}(D), \frac{1}{2}(C)\right], & \tau_{\mathrm{II}}(\phi) = C, \\
\tau_{\mathrm{I}}(DD) = D, & \tau_{\mathrm{II}}(DD) = \left[\frac{3}{4}(D), \frac{1}{4}(C)\right], \\
\tau_{\mathrm{I}}(DC) = \left[\frac{2}{3}(D), \frac{1}{3}(C)\right], & \tau_{\mathrm{II}}(DC) = \left[\frac{1}{2}(D), \frac{1}{2}(C)\right], \\
\tau_{\mathrm{I}}(CD) = \left[\frac{1}{4}(D), \frac{3}{4}(C)\right], & \tau_{\mathrm{II}}(CD) = C, \\
\tau_{\mathrm{I}}(CC) = C & \tau_{\mathrm{II}}(CC) = D.
\end{array}$$

Figure 13.4 Strategies for both players in the two-stage Prisoner's Dilemma ◀

Given the strategies $(\tau_i)_{i \in N}$ of the players, denote by $\tau = (\tau_1, \tau_2, \ldots, \tau_n)$ the vector of the players' strategies. Denote by $\tau_i(s_i)$ the probability that player i plays action s_i in the first stage, and by $\tau_i(s_i \mid s^1, \ldots, s^{t-1})$ the conditional probability that player i plays action s_i in stage t, given that the players have played $(s^1, \ldots, s^{t-1})$ in the first $t-1$ stages.

Example 13.1 (*Continued*) If the players play according to the strategies τ_{I} and τ_{II} that we defined in Figure 13.4 in the two-stage Prisoner's Dilemma, we can associate with every branch in the game tree the probability that it will be chosen in a play of the game. These probabilities are shown in Figure 13.5. The figure also shows, by each leaf of the game tree, the probability that the leaf will be arrived at if the players play strategies τ_{I} and τ_{II}.

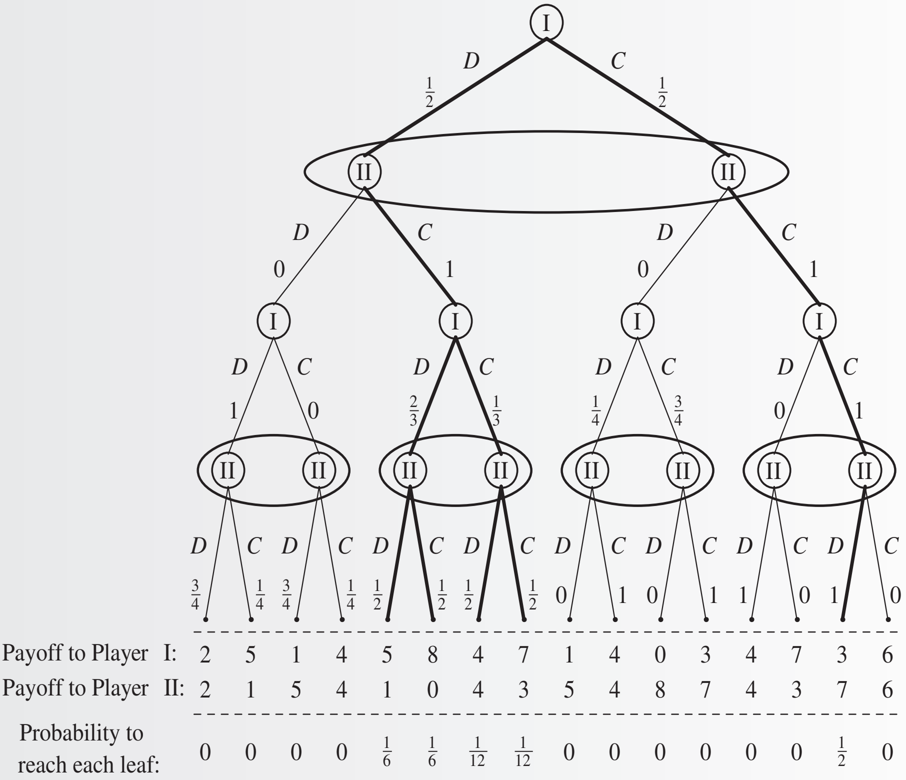

Figure 13.5 The probabilities attached to each play of the game, under the strategies $(\tau_{\mathrm{I}}, \tau_{\mathrm{II}})$ ◀

The collection of all the possible plays of the T-stage game is $S^T = H(T)$. As can be seen in Figure 13.5, every strategy vector τ naturally induces a probability measure $\mathbf{P}_\tau$ over $H(T)$. The probability of every play of the game $(s^1, s^2, \ldots, s^T)$ is the probability that if the players play according to strategy τ, the resulting play of the game will be this history. Formally, for every action vector $s^1 = (s_1^1, \ldots, s_n^1) \in S$, define

$$\mathbf{P}_\tau(s^1) = \tau_1\big(s_1^1\big) \times \tau_2\big(s_2^1\big) \times \cdots \times \tau_n\big(s_n^1\big). \tag{13.4}$$

This is the probability that the action vector played in the first stage is s^1, and it equals the product of the probability that every player i plays action s_i^1. More generally, for every t, $2 \leq t \leq T$, and every finite history$(s^1, s^2, \dots, s^t) \in S^t$, define by induction

$$\begin{aligned} \mathbf{P}_\tau(s^1, s^2, \dots, s^t) = \mathbf{P}_\tau(s^1, s^2, \dots, s^{t-1}) \times \ \tau_1\left(s_1^t \mid s^1, s^2, \dots, s^{t-1}\right) \\ \times \ \tau_2\left(s_2^t \mid s^1, s^2, \dots, s^{t-1}\right) \times \cdots \times \tau_n\left(s_n^t \mid s^1, s^2, \dots, s^{t-1}\right). \end{aligned}$$

This means that the probability that under τ the players play the action vector $s^1, s^2, \dots, s^t$ in the first t stages is the probability that the players play $s^1, s^2, \dots, s^{t-1}$ in the first $t-1$ stages, times the conditional probability that they play the action vector s^t in stage t, given that they played $s^1, s^2, \dots, s^{t-1}$ in the first $t-1$ stages. This formula for $\mathbf{P}_\tau$ expresses the fact that the mixed action that a player implements in any given stage can depend on the actions that he or other players played in previous stages, but the random choices of the players made simultaneously in each stage are independent of each other. The case in which there may be correlation between the actions chosen by the players was addressed in Chapter 8, where we studied the concept of correlated equilibrium.

13.3.2 Payoffs and equilibria

In repeated games, the players receive a payoff in every stage of the game. Denote the payoff received by player i in stage t by u_i^t, and denote the vector of payoffs to the players in stage t by $u^t = (u_1^t, \dots, u_n^t)$. Then, during the course of a play of the game, player i receives the sequence of payoffs $(u_i^1, u_i^2, \dots, u_i^T)$. We assume that every player seeks to maximize the sum total of these payoffs or, equivalently, seeks to maximize the average of these payoffs.

As previously noted, every strategy vector τ induces a probability measure $\mathbf{P}_\tau$ over $H(T)$. Denote the corresponding expectation operator by $\mathbf{E}_\tau$; i.e., for every function $f : H(T) \to \mathbb{R}$, the expectation of f under $\mathbf{P}_\tau$ is denoted by $\mathbf{E}_\tau[f]$:

$$\mathbf{E}_\tau[f] = \sum_{(s^1,\dots,s^T)\in H(T)} \mathbf{P}_\tau(s^1, \dots, s^T) f(s^1, \dots, s^T). \tag{13.5}$$

Player i's expected payoff in stage t, under the strategy vector τ, is $\mathbf{E}_\tau[u_i^t]$. Denote player i's average expected payoff in the first T stages under strategy vector τ by

$$\gamma_i^T(\tau) := \mathbf{E}_\tau\left[\frac{1}{T}\sum_{t=1}^{T} u_i^t\right] = \frac{1}{T}\sum_{t=1}^{T} \mathbf{E}_\tau\left(u_i^t\right). \tag{13.6}$$

Example 13.1 (*Continued*) Figure 13.5 provides the probability to every play of the game under the strategy pair $(\tau_{\mathrm{I}}, \tau_{\mathrm{II}})$. The table in Figure 13.6 presents the plays of the game that are obtained with positive probability in the left column, the probability that each play is obtained in the middle column, and the payoff to the players, under that play of the game, in the right column. Each play of the game is written from left to right, with the actions implemented by the players in the first stage appearing first, followed by the actions implemented by the players in the second stage. Player I's action appears to the left of Player II's action.

Play of the Game	Probability	Payoff
$(D, C), (D, D)$	$\frac{1}{6}$	$(5, 1)$
$(D, C), (D, C)$	$\frac{1}{6}$	$(8, 0)$
$(D, C), (C, D)$	$\frac{1}{12}$	$(4, 4)$
$(D, C), (C, C)$	$\frac{1}{12}$	$(7, 3)$
$(C, C), (C, D)$	$\frac{1}{2}$	$(3, 7)$

Figure 13.6 The probability of every play of the game, and the corresponding payoff, under the strategy pair $(\tau_{\mathrm{I}}, \tau_{\mathrm{II}})$

It follows that the expected payoff of the two players is

$$\tfrac{1}{6} \times (5, 1) + \tfrac{1}{6} \times (8, 0) + \tfrac{1}{12} \times (4, 4) + \tfrac{1}{12} \times (7, 3) + \tfrac{1}{2} \times (3, 7) = \left(4\tfrac{7}{12}, 4\tfrac{1}{4}\right). \tag{13.7}$$

◀

Definition 13.5 *Let* $\Gamma = (N, (S_i)_{i\in N}, (u_i)_{i\in N})$ *be a base game. The* T-stage game Γ_T *corresponding to* Γ *is the game* $\Gamma_T = (N, (\mathcal{B}_i^T)_{i\in N}, (\gamma_i^T)_{i\in N})$.

The strategy vector $\tau^* = (\tau_1^*, \ldots, \tau_n^*)$ is a (Nash) equilibrium of Γ_T if for each player $i \in \mathbb{N}$, and each strategy $\tau_i \in \mathcal{B}_i^T$,

$$\gamma_i^T(\tau^*) \geq \gamma_i^T(\tau_i, \tau_{-i}^*). \tag{13.8}$$

The vector $\gamma^T(\tau^*)$ is called an *equilibrium payoff* of the repeated game Γ_T.

The following theorem states that a strategy vector at which in each stage the players play a one-stage equilibrium is an equilibrium of the T-stage game.

Theorem 13.6 *Let* $\Gamma = (N, (S_i)_{i\in N}, (u_i)_{i\in N})$ *be a base game, and let* Γ_T *be its corresponding repeated* T*-stage game. Let* $\sigma^1, \sigma^2, \ldots, \sigma^T$ *be equilibria of* Γ *(not necessarily different equilibria). Then the strategy vector* τ^* *in* Γ_T*, at which in each stage* t*,* $1 \leq t \leq T$*, every player* $i \in N$ *plays the mixed action* σ_i^t*, is an equilibrium.*

Proof: The strategy vector τ^* is an equilibrium, because neither player can profit by deviating. No player can profit in a stage in which he deviates from equilibrium, because by definition in such a stage the players implement an equilibrium of the base game. In addition, his deviation in any stage cannot influence the future actions of the other players, because they are playing according to a strategy that depends only on the stage t, not on the history h^t.

Formally, let $i \in N$ be a player, and let τ_i be any strategy of player i in Γ_T. We will show that $\gamma_i^T(\tau_i, \tau_{-i}^*) \leq \gamma_i^T(\tau^*)$; i.e., player i does not profit by deviating from τ_i^* to τ_i.

For each t, $1 \leq t \leq T$, the mixed action vector σ^t is an equilibrium of Γ. Therefore, for each history $h^{t-1} \in H(t-1)$,

$$u_i(\sigma^t) \geq u_i\big(\tau_i(h^{t-1}), \sigma_{-i}^t\big). \tag{13.9}$$

This implies that

$$\mathbf{E}_{\tau_i,\tau^*_{-i}}\left[u^t_i\right] = \sum_{h^{t-1}\in H(t-1)} \mathbf{P}_{\tau_i,\tau^*_{-i}}(h^{t-1}) u_i(\tau_i(h^{t-1}), \tau^*_{-i}(h^{t-1})) \tag{13.10}$$

$$= \sum_{h^{t-1}\in H(t-1)} \mathbf{P}_{\tau_i,\tau^*_{-i}}(h^{t-1}) u_i\left(\tau_i(h^{t-1}), \sigma^t_{-i}\right) \tag{13.11}$$

$$\leq \sum_{h^{t-1}\in H(t-1)} \mathbf{P}_{\tau_i,\tau^*_{-i}}(h^{t-1}) u_i(\sigma^t) \tag{13.12}$$

$$= u_i(\sigma^t) \sum_{h^{t-1}\in H(t-1)} \mathbf{P}_{\tau_i,\tau^*_{-i}}(h^{t-1}) = u_i(\sigma^t). \tag{13.13}$$

The last equality follows from the fact that the sum total of the probabilities of all $(t-1)$-stage histories is 1, and therefore $\mathbf{E}_{\tau_i,\tau^*_{-i}}[u^t_i] \leq u_i(\sigma^t)$. Averaging over the T stages of the game shows that $\gamma^T_i(\tau_i, \tau^*_{-i}) \leq \gamma^T_i(\tau^*)$, which is what we wanted to show. Since $\gamma^T_i(\tau_i, \tau^*_{-i}) \leq \gamma^T_i(\tau^*)$ for every strategy τ_i of player i, and for every player i, we deduce that τ^* is an equilibrium. □

By repeating the same equilibrium in every stage, we get the following corollary.

Corollary 13.7 *Let Γ be a base game, and let Γ_T be the corresponding repeated T-stage game. Every equilibrium payoff of Γ is also an equilibrium payoff of Γ_T.*

13.3.3 The minmax value

Recall that U_i is the multilinear extension of u_i (Equation (5.9), page 147). The minmax value of player i in the base game Γ is (Equation (4.51), page 113):

$$\overline{v}_i = \min_{\sigma_{-i}\in\times_{j\neq i}\Sigma_j} \max_{\sigma_i\in\Sigma_i} U_i(\sigma_i, \sigma_{-i}). \tag{13.14}$$

This is the value that the players $N \setminus \{i\}$ cannot prevent player i from attaining: for any vector of mixed actions σ_{-i} they implement, player i can receive at least $\max_{\sigma_i\in\Sigma_i} U_i(\sigma_i, \sigma_{-i})$, which is at least $\overline{v}_i$. Every mixed strategy vector σ_{-i} satisfying

$$\overline{v}_i = \max_{\sigma_i\in\Sigma_i} U_i(\sigma_i, \sigma_{-i}) \tag{13.15}$$

is called a *punishment strategy vector* against player i, because if the players $N \setminus \{i\}$ play σ_{-i}, they guarantee that player i's average payoff will not exceed $\overline{v}_i$. Similarly to what we saw in Equation (5.25) (page 151), for every mixed action vector $\sigma_{-i} \in \Sigma_{-i}$ there exists a pure action $s'_i \in S_i$ of player i satisfying $U_i(s'_i, \sigma_{-i}) \geq \overline{v}_i$ (why?).

The next theorem states that at every equilibrium of the repeated game, the payoff to each player i is at least $\overline{v}_i$. The discussion above and the proof of the theorem imply that the minmax value of each player i in the T-stage game is $\overline{v}_i$ (Exercise 13.8).

Theorem 13.8 *Let τ^* be an equilibrium of Γ_T. Then $\gamma^T_i(\tau^*) \geq \overline{v}_i$ for each player $i \in \mathbb{N}$.*

Proof: We will show that for every strategy vector τ (not necessarily an equilibrium vector) there exists a strategy τ^*_i of player i (which depends on τ_{-i}) satisfying $\gamma^T_i(\tau^*_i, \tau_{-i}) \geq \overline{v}_i$.

It follows, in particular, that if τ is an equilibrium, then

$$\gamma_i^T(\tau) \geq \gamma_i^T(\tau_i^*, \tau_{-i}) \geq \overline{v}_i, \tag{13.16}$$

which is what the theorem claims. We now construct such a strategy τ_i^* explicitly, for any given τ_{-i}. Recall that when τ is a strategy vector, $\tau_j(h)$ is the mixed action that player j plays after history h, and $\tau_{-i}(h) = (\tau_j(h))_{j \neq i}$ is the mixed action vector that the players $N \setminus \{i\}$ play after history h. As previously noted, for every history $h \in \bigcup_{t=0}^{T-1} H(t)$ there is an action $s_i'(h) \in S_i$ such that $U_i(s_i'(h), \tau_{-i}(h)) \geq \overline{v}_i$. Let τ_i^* be a strategy of player i under which, after every history h, he plays the action $s_i'(h)$. Then for every $t \in \{1, 2, \ldots, T\}$,

$$\mathbf{E}_{\tau_i^*,\tau_{-i}}\left[u_i^t\right] = \sum_{h^{t-1}\in H(t-1)} \mathbf{P}_{\tau_i^*,\tau_{-i}}(h^{t-1}) u_i(\tau_i^*(h^{t-1}), \tau_{-i}(h^{t-1})) \tag{13.17}$$

$$= \sum_{h^{t-1}\in H(t-1)} \mathbf{P}_{\tau_i^*,\tau_{-i}}(h^{t-1}) u_i(s_i'(h^{t-1}), \tau_{-i}(h^{t-1})) \tag{13.18}$$

$$\geq \sum_{h^{t-1}\in H(t-1)} \mathbf{P}_{\tau_i^*,\tau_{-i}}(h^{t-1}) \overline{v}_i = \overline{v}_i. \tag{13.19}$$

The last equality follows from the fact that the sum total of the probabilities of all the possible histories at time period t is 1. In words, the expected payoff in stage t is at least $\overline{v}_i$. By averaging over the T stages of the game, we conclude that the expected average of the payoffs is at least $\overline{v}_i$:

$$\gamma_i^T(\tau_i^*, \tau_{-i}) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{E}_{\tau_i^*,\tau_{-i}}[u_i^t] \geq \frac{1}{T}\sum_{t=1}^{T} \overline{v}_i = \overline{v}_i, \tag{13.20}$$

which is what we wanted to show. □

Define a set of payoff vectors V by

$$V := \left\{x \in \mathbb{R}^N : x_i \geq \overline{v}_i \text{ for each player } i \in N\right\}. \tag{13.21}$$

This is the set of payoff vectors at which every player receives at least his minmax value. The set is called the set of *individually rational payoffs*. Theorem 13.8 implies that the set of equilibrium payoffs is contained in V.

13.4 Characterization of the set of equilibrium payoffs of the *T*-stage repeated game

For every set of vectors $\{x_1, \ldots, x_K\}$ in $\mathbb{R}^N$, denote by $\text{conv}\{x_1, \ldots, x_K\}$ the smallest convex set that contains $\{x_1, \ldots, x_K\}$.

The players play some action vector s in S in each stage; hence the payoff vector in each stage is one of the vectors $\{u(s), s \in S\}$. In particular, the average payoff of the players, which is equal to $\frac{1}{T}\sum_{t=1}^{T} u(s^t)$, is necessarily located in the convex hull of these vectors (because it is a weighted average of the vectors in this set), which we denote by F:

$$F := \text{conv}\{u(s), s \in S\}. \tag{13.22}$$

This set is called the *set of feasible payoffs*. We thus have $\gamma^T(\tau) \in F$ for every strategy vector τ.

Using the last remark, and Theorem 13.8, we deduce that the set of equilibrium payoffs is contained in the set $F \cap V$ of feasible and individually rational payoff vectors. As we now show, if the base game satisfies a certain technical condition, then for every feasible and individually rational payoff vector x there exists an equilibrium payoff vector of the T-stage game that is close to it, for sufficiently large T. The technical condition that is needed here is that, for every player i, it is possible to find an equilibrium of the base game at which the payoff to player i is strictly greater than his minmax value.

Theorem 13.9 (The Folk Theorem[3]) *Suppose that for every player $i \in \mathbb{N}$ there exists an equilibrium $\beta(i)$ in the base game $\Gamma = (N, (S_i)_{i\in N}, (u_i)_{i\in N})$ satisfying $u_i(\beta(i)) > \overline{v}_i$. Then for every $\varepsilon > 0$ there exists $T_0 \in \mathbb{N}$ such that for every $T \geq T_0$, and every feasible and individually rational payoff vector $x \in F \cap V$, there exists an equilibrium τ^* of the T-stage game Γ_T whose corresponding payoff is ε-close to x (in the maximum norm[4]):*

$$\|\gamma^T(\tau^*) - x\|_\infty < \varepsilon. \tag{13.23}$$

Under every equilibrium β of the base game, $u_i(\beta) \geq \overline{v}_i$ for every player i (as implied by Theorem 13.8 for $T = 1$). The condition of the theorem requires furthermore that, for every player i, there exist an equilibrium at which that inequality is a strict inequality.

Remark 13.10 *One can choose the minimal length T_0 in Theorem 13.9 to be independent of x. To see this, note that since $F \cap V$ is a compact set, given ε there exists a finite set $x^1, x^2, \ldots, x^J$ of vectors in $F \cap V$ such that the distance between each vector $x \in F$ and at least one of the vectors $x^1, x^2, \ldots, x^J$ is below $\frac{\varepsilon}{2}$:*

$$\max_{x\in F\cap V} \min_{1\leq j\leq J} \|x - x^j\|_\infty \leq \frac{\varepsilon}{2}. \tag{13.24}$$

Denote by $T_0(x^j, \frac{\varepsilon}{2})$ the size of T_0 in Theorem 13.9 corresponding to x^j and $\frac{\varepsilon}{2}$. Let $x \in F \cap V$, and let $j_0 \in \{1, 2, \ldots, J\}$ be an index satisfying $\|x - x^{j_0}\|_\infty \leq \frac{\varepsilon}{2}$. By the triangle inequality, every equilibrium τ of the T-stage repeated game satisfying $\|\gamma^T(\tau) - x^{j_0}\|_\infty \leq \frac{\varepsilon}{2}$ also satisfies $\|\gamma^T(\tau) - x\|_\infty \leq \varepsilon$. It follows that the statement of Theorem 13.9 holds for x and ε with $T_0 := \max_{1\leq j\leq J} T_0(x^j, \frac{\varepsilon}{2})$, and this T_0 is independent of x. ◆

13.4.1 Proof of the Folk Theorem: example

Before we prove the theorem, we present an example that illustrates the proof. Consider the two-player game in Figure 13.7 (this is the game of Chicken; see Example 8.3 on page 303).

The minmax value of both players is 2. The punishment strategy against Player I is R, and the punishment strategy against Player II is B. The game has two equilibria in pure

---

[3] The name of the Folk Theorem is borrowed from the analogous theorem (see Theorem 13.17) for infinitely repeated games, which was well known in the scientific community for many years, despite the fact that it was not formally published in any journal article, and hence it was called a "folk theorem." The theorem is now usually ascribed to Aumann and Shapley [1994]. The Folk Theorem for finite games, Theorem 13.9, was proved by Benoit and Krishna [1985].

[4] The maximum norm over $\mathbb{R}^n$ is defined as follows: $\|x\|_\infty = \max_{i=1,2,\ldots,n} |x_i|$ for each vector $x \in \mathbb{R}^n$.

		Player II	
		L	R
Player I	T	6, 6	2, 7
	B	7, 2	0, 0

Figure 13.7 The payoff matrix of the game of Chicken

strategies, (T, R) and (B, L), with payoffs (2, 7) and (7, 2) respectively (we will not use the equilibrium in mixed strategies). If we denote

$$\beta(\mathrm{I}) = (B, L), \qquad \beta(\mathrm{II}) = (T, R), \tag{13.25}$$

we deduce that the condition of Theorem 13.9 holds (because $u_i(\beta(i)) = 7 > 2 = \overline{v}_i$ for $i \in \{\mathrm{I}, \mathrm{II}\}$).

The payoff vector (3, 3) is in F, since $(3, 3) = \frac{1}{2}(0, 0) + \frac{1}{2}(6, 6)$. It is also in V, because both of its coordinates are greater than or equal to 2, which is the minmax value of both players. It is therefore in $F \cap V$. We will now construct an equilibrium of the 100-stage game, whose average payoff is close to (3, 3).

If the players play (T, L) in odd-numbered stages (yielding the payoff (6, 6) in every odd-numbered stage) and play (B, R) in even-numbered stages (yielding the payoff (0, 0) in every even-numbered stage), the average payoff is (3, 3). This does not yet constitute an equilibrium, because every player can profit by deviating at every stage. Because this sequence of actions is deterministic, any deviation from it is immediately detected, and the other player can then implement the punishment strategy. The punishment strategy guarantees that the deviating player receives at most 2 in every stage after the deviation, which is less than the average of 3 that he can receive if he avoids deviating.

Because the repeated game in this case is finite, a threat to implement a punishment strategy is effective only if there are sufficiently many stages left to guarantee that the loss imposed on a deviating player is greater than the reward he stands to gain by deviating. If, for example, a player deviates in the last stage, he cannot be punished because there are no more stages, and he therefore stands to gain by such a deviation. This detail has to be taken into consideration in constructing an equilibrium.

We now describe a strategy vector defined by a basic plan of action and a punishment strategy. The basic plan of action is depicted in Figure 13.8, and consists of 49 cycles, each comprised of two stages, along with a tail-end that is also comprised of two stages.

In the first 98 stages, the players alternately play the action vectors (T, L) and (B, R), thereby guaranteeing that the average payoffs in these stages is (3, 3), with the average payoff in all 100 stages close to (3, 3). In these stages, they play according to a deterministic plan of action; hence if one of them deviates from this plan, the other immediately takes note of this deviation. Once one player deviates at a certain stage, the other player implements the punishment strategy against the deviator, from the next stage on: if Player II deviates, Player I plays B from the next stage to the end of the play of the game. If

Player I's actions	T	B	T	B	•	•	•	T	B	B	T
Player II's actions	L	R	L	R	•	•	•	L	R	L	R
Stage	1	2	3	4	•	•	•	97	98	99	100
Player I's payoff	6	0	6	0	•	•	•	6	0	7	2
Player II's payoff	6	0	6	0	•	•	•	6	0	2	7

Figure 13.8 An equilibrium in the 100-stage game of Chicken

Player I deviates, Player II plays R from the next stage to the end of the play of the game. In the last two stages of the basic plan of action, the players play the pure strategy equilibria $\beta(\mathrm{I})$ and $\beta(\mathrm{II})$ (in that order).

We now show that this strategy vector is an equilibrium yielding an average payoff that is close to (3, 3). Indeed, if the players follow this strategy vector, the average payoff is

$$\tfrac{49}{100}(6,6)+\tfrac{49}{100}(0,0)+\tfrac{1}{100}(7,2)+\tfrac{1}{100}(2,7)=(3.03,3.03), \tag{13.26}$$

which is close to (3, 3).

We next turn to ascertaining that Player I cannot gain by deviating (ascertaining that Player II cannot gain by deviating is conducted in a similar way). In each of the last two stages (the tail-end of the action plan), the two players play an equilibrium of the base game, and therefore Player I cannot gain by deviating in those stages. Suppose, therefore, that Player I deviated during one of the first 98 stages. In the cycle at which he deviates for the first time, he can gain at most 3, relative to the payoff he would receive at that cycle by following the basic action plan. To see this, note that if he deviates in the second stage of the cycle (playing T instead of B), he gains 2 at that stage. If he deviates in the first stage of the cycle (playing B instead of T), he gains 1 at that stage, and if he then plays T instead of B in the second stage of the cycle he gains 2 at that stage, and in total he gains 3 at that cycle ($7+2$ instead of $6+0$ according to the basic plan). In each of the following cycles he loses (because he receives at most 2 in every stage of the cycle, instead of receiving 6 in the first stage and 0 in the second stage of the cycle, as he would receive under the basic plan of action). Finally, at stage 100 he loses 5: he will receive at most 2 rather than the 7 that he receives in the basic plan of action. In sum total, the deviation leads to a loss of at least $5-3=2$, relative to the payoff he would receive by following the basic action plan, and therefore Player I cannot gain by deviating.

In the construction depicted here, we have split the stages into cycles of length 2, because the payoff (3, 3) is the average of two payoff vectors of the matrix. If we had wanted to construct an equilibrium with a payoff that is, say, close to $(3\frac{1}{2},4\frac{3}{4})$ (which is also in $F\cap V$), then, since $(3\frac{1}{2},4\frac{3}{4})=\frac{1}{4}(0,0)+\frac{1}{2}(6,6)+\frac{1}{4}(2,7)$, we would have constructed an equilibrium using cycles of length 4: except for the last stages, the players would repeatedly play the action vectors

$$(B,R),(T,L),(T,L),(T,R). \tag{13.27}$$

We can mimic the construction above whenever the target payoff can be obtained as the weighted average of the payoff vectors in the matrix, with rational weights.

Since the target payoff is in F, it can always be obtained as a weighted average of payoffs. If the weights are irrational, we need to approximate them using rational weights.

The role of the tail-end (the last two stages in the above example) is to guarantee that a deviating player loses. During the course of the tail-end, the players cyclically play the equilibria $\beta(1), \ldots, \beta(n)$. The expected payoff of each player i under each of these equilibria is greater than or equal to $\overline{v}_i$ (because they are equilibria) and under $\beta(i)$ it is strictly greater than $\overline{v}_i$. That is why, if the other players punish player i by reducing his payoff to $\overline{v}_i$, he loses in the tail-end. The tail-end needs to be sufficiently long for the total loss to be greater than the maximal gain that a player can obtain by deviating. On the other hand, the tail-end needs to be sufficiently short, relative to the length of the game, for the overall payoff to be close to the target payoff (which is the average payoff in a single cycle).

In the formulation of the Folk Theorem, the equilibrium payoff does not equal the target payoff x; the best we can do is obtain a payoff that is close to it. This stems from two reasons:

1. The existence of the tail-end, in which the payoff is not the target payoff.
2. It may be the case that x cannot be expressed as the weighted average of payoff vectors of the matrix using rational weights, which then requires approximating these weights using rational weights.

13.4.2 Detailed proof of the Folk Theorem

We will now generalize the construction in the example of the previous section to all repeated games. For every real number c, denote by $\lceil c \rceil$ the least integer that is greater than or equal to c, and by $\lfloor c \rfloor$ the greatest integer that is less than or equal to c. Recall that $M = \max_{i\in N} \max_{s\in S} |u_i(s)|$ is the maximal payoff of the game (in absolute value).

Step 1: Determining the cycle length.
We first show that every vector in F can be approximated by a weighted average of the vectors $(u(s))_{s\in S}$, with rational weights sharing the same denominator. The proof of the following theorem is left to the reader (Exercise 13.13).

Theorem 13.11 *For every $K \in \mathbb{N}$ and every vector $x \in F$ there are nonnegative integers $(k_s)_{s\in S}$ summing to K satisfying*

$$\left\| \sum_{s\in S} \frac{k_s}{K} u(s) - x \right\|_\infty \leq \frac{M \times |S|}{K}. \tag{13.28}$$

For $\varepsilon > 0$ and $x \in F \cap V$, let K be a natural number satisfying $K \geq \frac{2M\times|S|}{\varepsilon}$ and let $(k_s)_{s\in S}$ be nonnegative integers summing to K satisfying Equation (13.28). If the players implement cycles of length K, and in each cycle they play each action vector $s \in S$ exactly k_s times, then the average payoff over the course of the cycle is $\sum_{s\in S} \frac{k_s}{K} u(s)$, and the distance between this average payoff and x is at most $\frac{M\times|S|}{K}$.

Step 2: Defining the strategy vector τ^*.
We next define a strategy vector τ^* of the T-stage game, which depends on two variables, R and L, to be defined later. The T stages of the game are divided into R cycles of length K and a tail of length L:

$$T = RK + L. \tag{13.29}$$

These variables will be set in such a way that the following two properties are satisfied: R will be sufficiently large for the average payoff according to τ^* to be close to x, and L will be sufficiently large for τ^* to be an equilibrium. In each cycle, the players play every action vector $s \in S$ exactly k_s times. In the tail-end, the players cycle through the equilibria $\beta(1), \ldots, \beta(n)$. In other words, each player j plays the mixed action $\beta_j(1)$ in the first stage, and in stages $n+1$, $2n+1$, etc., of the tail-end; he plays the mixed action $\beta_j(2)$ in the second stage, and in stages $n+2$, $2n+2$, etc., of the tail-end, and so on.

The basic plan that we have defined for the first RK stages is deterministic: the players do not choose their actions randomly in these stages. It follows that if a player deviates from the basic plan in one of the first RK stages, this deviation is detected by the other players. In this case, from the next stage on, the other players punish the deviator: at every subsequent stage they implement a punishment strategy vector against the deviator. If a player deviates for the first time in one of the L final stages, the other players do not punish him, and instead continue cycling through the equilibria $\{\beta(i)\}_{i\in N}$.

Step 3: The constraints on R and L needed to ensure that the distance between the average payoff under τ^* and x is at most ε.
Suppose that the players implement the strategy vector τ^*. Given the choice of $(k_s)_{s\in S}$, the distance between the average payoff in every cycle of length K and x is at most $\frac{M\times|S|}{K}$. This also holds true for any integer number of repetitions of the cycle. By the choice of K, one has $\frac{M\times|S|}{K} \leq \frac{\varepsilon}{2}$, and hence the distance between the average payoff in the first RK stages and x is at most $\frac{\varepsilon}{2}$. If the length of the tail-end L is small relative to RK, the average payoff in the entire game will be close to x. We will ascertain that if

$$L \leq \frac{KR\varepsilon}{4M}, \tag{13.30}$$

then the distance between the average payoff in the entire game and x is at most ε. Indeed, the distance between the average payoff in the first RK stages and x is at most $\frac{\varepsilon}{2}$, and the distance between the average payoff in the last L stages and x is at most $2M$. Therefore the average payoff in the entire game is within ε of x, as long as

$$\frac{RK\frac{\varepsilon}{2} + 2ML}{T} \leq \varepsilon. \tag{13.31}$$

Since $T = RK + L > RK$, it suffices to require that

$$\frac{RK\frac{\varepsilon}{2} + 2ML}{RK} \leq \varepsilon, \tag{13.32}$$

and this inequality is equivalent to Equation (13.30).

Step 4: τ^* is an equilibrium.
Suppose that player i first deviates from the basic plan at stage t_0. We will ascertain here that his average payoff cannot increase by such a deviation.

Suppose first that t_0 is in the tail-end: $t_0 > RK$. Since throughout the tail the players play an equilibrium of the base game at every stage, player i cannot increase his average payoff by such a deviation.

Suppose next that $t_0 \leq RK$. Then player i's deviation triggers a punishment strategy against him from stage $t_0 + 1$. It follows that from stage $t_0 + 1$ player i's payoff at each stage is at most his minmax value $\overline{v}_i$. If $L \geq n$, by the condition that $u_i(\beta(i)) > \overline{v}_i$ we deduce that at each n consecutive stages in the tail-end, player i loses by the deviation at least $u_i(\beta(i)) - \overline{v}_i$, relative to his payoff at the equilibrium strategy. Denote $\delta_i = u_i(\beta(i)) - \overline{v}_i > 0$, and $\delta = \min_{i \in N} \delta_i > 0$.

The maximal profit that player i can gain by deviating up to stage RK is $2KM$: because the payoffs are between $-M$ and M, player i can gain at most $2M$ by deviating in any single stage; hence in a cycle in which he deviates, a player can gain[5] at most $2KM$. The player cannot gain in any of the subsequent cycles, because the average payoff in a cycle under the equilibrium strategy is x, while if a player deviates, he receives at most $\overline{v}_i$, while $\overline{v}_i \leq x_i$.

For a punishment to be effective, we need to require that the tail-end be sufficiently long to ensure that the losses at the tail-end exceed the possible gains in the cycle in which the deviation occurred:

$$\delta \left\lfloor \frac{L}{n} \right\rfloor > 2KM. \tag{13.33}$$

In this calculation, we have rounded down L/n. In every n stages of the tail-end, every player is punished only once. If L is not divisible by n, some of the players are punished $\lfloor \frac{L}{n} \rfloor$ times, and some are punished $\lceil \frac{L}{n} \rceil$ times.

Equation (13.33) gives us the required minimal length of the tail-end

$$L > n\left(1 + \frac{2KM}{\delta}\right). \tag{13.34}$$

The length of the tail-end, L, cannot be constant for all T, because $T - L$ needs to be divisible by K. It suffices to use tail-ends whose length is at least $n\left(1 + \frac{2KM}{\delta}\right)$, and at most $n\left(1 + \frac{2KM}{\delta}\right) + K$.

Step 5: Establishing T_0.
The length of the game, T, satisfies $T = RK + L$. From Equation (13.30), we need to require that $R \geq \frac{4ML}{K\varepsilon}$, i.e., $T = RK + L \geq L\left(1 + \frac{4M}{\varepsilon}\right)$. This, along with Equation (13.34), implies that the length of the game must satisfy

$$T > n\left(1 + \frac{2KM}{\delta}\right)\left(1 + \frac{4M}{\varepsilon}\right). \tag{13.35}$$

---

[5] If a player deviates at any stage, from the next stage on his one-stage expected payoff is at most his minmax value, but it is possible that in the basic plan during the cycle there may be stages in which his payoff is less than his minmax value. For example, in the equilibrium constructed in the example in Section 13.4.1 (page 531), in the even stages the payoff to each player is 0, while the minmax value of each player is 2. It is therefore possible for a player to gain at more than one stage by deviating.

We can therefore set T_0 to be the value of the right-hand side of Equation (13.35). This concludes the proof of Theorem 13.9.

Remark 13.12 *As mentioned above, the only equilibrium payoff in the finitely repeated Prisoner's Dilemma is* $(1, 1)$*. This does not contradict Theorem 13.9, because the conditions of the theorem do not hold in this case: the only equilibrium of the one-stage Prisoner's Dilemma is* (D, D)*, and the payoff to both players at this equilibrium is* 1*, which is the minmax value of both players. The proof of the uniqueness of the equilibrium payoff in the* T*-stage Prisoner's Dilemma is based on the existence of a last stage in the game. In the next section we will study repeated games of infinite length, and show that in that case, the repeated Prisoner's Dilemma has more than one equilibrium payoff.* ◆

13.5 Infinitely repeated games

As noted above, the strategy vector constructed in the previous section is highly dependent on the length of the game: it cannot be implemented unless the players know the length of the game. However, it is often the case that the length of a repeated game is not known ahead of time. For example, the owner of a tennis-goods shop does not know if or when he will sell his shop, tennis players do not know when they will stop playing tennis, nor if or when they will move to another town. Infinitely repeated games can serve to model finite but extremely long repeated games, in which (a) the number of stages is unknown, (b) the players ascribe no importance to the last stage of the game, or (c) at every stage the players believe that the game will continue for several more stages.

In this section, we will present a model of infinitely repeated games, and characterize the set of equilibria of such games. The definitions in this section are analogous to the definitions in the section on T-stage games. As the next example shows, extending games to an infinite number of repetitions leads to new equilibrium payoffs: payoff vectors that cannot be obtained as limits of sequences of equilibrium payoffs in finite games whose lengths increase to infinity.

Example 13.1 (*Continued*) Recall the repeated Prisoner's Dilemma, given by the payoff matrix in Figure 13.9.

		Player II	
		D	C
Player I	D	1, 1	4, 0
	C	0, 4	3, 3

Figure 13.9 The Prisoner's Dilemma

Consider the repeated Prisoner's Dilemma in the case where the players repeat playing the basic game ad infinitum. In this case, every player receives an infinite sequence of payoffs: one payoff per stage of the game. We will assume that every player strives to maximize the limit of the average payoff he receives. Certain technical issues, such as what happens when the limit of the average payoffs does not exist, will be temporarily ignored (we will consider this issue later in this chapter). Whereas the only equilibrium payoff of the T-stage repeated game is $(1, 1)$, in the infinitely repeated game there are additional equilibrium payoffs.

Let us look, for example, at following strategy: in the first stage play C. In every subsequent stage, if the other player chose C in every stage since the game started, choose C in the current stage; otherwise choose D. This is an unforgiving strategy that is called the Grim-Trigger Strategy: as long as the opponent cooperates, you also cooperate, but if he fails to cooperate once, defect forever from that point on in the game. When both players implement this strategy, no player has an incentive to deviate. To see why, note that at this strategy vector, every player receives 3 in each stage; hence every player's average payoff is 3. If a player deviates at some stage and plays D, he receives 4 in that stage, instead of 3, but from that stage on, the other player plays D, and then the most that the deviating player receives in every stage is 1. In particular, the limit of the average payoff of the deviating player is at most 1. Thus, the payoff vector $(3, 3)$ is an equilibrium payoff of the infinitely repeated game, despite not being an equilibrium payoff of the T-stage game. ◀

Definition 13.13 *A* behavior strategy *for player i (in the infinitely repeated game) is a function mapping every finite history to a mixed action:*

$$\tau_i : \bigcup_{t=0}^{\infty} H(t) \to \Sigma_i. \qquad (13.36)$$

The collection of all of player i's strategies in the infinitely repeated game is denoted by $\mathcal{B}_i^\infty$.

Remark 13.14 *If $\tau_i(h) \in S_i$ for every finite history $h \in \bigcup_{t=1}^{\infty} H(t)$, then the strategy τ_i is a* pure strategy. *Note that even when the sets of actions of the players are finite, the set of pure strategies available to them has the cardinality of the continuum.* ◆

Denote by $H(\infty)$ the collection of all possible plays of the infinitely repeated game:

$$H(\infty) = S^{\mathbb{N}}. \qquad (13.37)$$

An element of this set is an infinite sequence $(s^1, s^2, \ldots)$ of action vectors, where $s^t = (s_i^t)_{i\in N}$ is the action vector of the players in stage t. The results in Section 6.4 (page 238) show that every vector of behavior strategies $\tau = (\tau_i)_{i\in N}$ induces a probability distribution $\mathbf{P}_\tau$ over the set $H(\infty)$ (which, together with the σ-algebra of cylinder sets forms a measurable space; see Section 6.4 for the definitions of these notions). We denote by $\mathbf{E}_\tau$ the expectation operator that corresponds to the probability distribution $\mathbf{P}_\tau$.

To define an infinitely repeated game, we need to define, in addition to the sets of strategies of the players, their payoff functions. One way to try doing this is by taking the limit of Equation (13.6) as T goes to infinity, but this limit may not necessarily exist. In this section, we will define infinitely repeated games, and equilibria in infinitely repeated games, without explicitly defining payoff functions in such games. In the next section we will define discounted payoff functions for infinitely repeated games, and study the corresponding equilibrium notion, which turns out to be different from the equilibrium concept presented in this section.

Definition 13.15 *Let* $\Gamma = (N, (S_i)_{i \in N}, (u_i)_{i \in N})$ *be a game in strategic form. The* infinitely repeated game Γ_∞ *corresponding to* Γ *is the game whose set of players is* N*, and each player* i*'s set of strategies is* $\mathcal{B}_i^\infty$.

For every stage $t \in \mathbb{N}$, and every player $i \in N$, denote by u_i^t player i's payoff in stage t. We next define the concept of the equilibrium of Γ_∞.

Definition 13.16 *A strategy vector* τ^* *is an* equilibrium *(of the infinitely repeated game* Γ_∞*), with a corresponding* equilibrium payoff $x \in \mathbb{R}^N$*, if with probability* 1 *according to* $\mathbf{P}_{\tau^*}$*, for each player* $i \in N$*, the limit*

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u_i^t \tag{13.38}$$

exists, and for each strategy τ_i *of player* i*,*

$$\mathbf{E}_{\tau^*}\left(\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u_i^t\right) = x_i \geq \mathbf{E}_{\tau_i, \tau_{-i}^*}\left(\limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u_i^t\right). \tag{13.39}$$

In words, a strategy vector τ^* is an equilibrium, with payoff x if (a) the average payoff under τ^* converges, (b) the expectation of its limit is x, and (c) no player can profit by deviating. Since there is no guarantee that every deviation leads to a well-defined limit of the average payoffs, we require that the expectation of the limit superior of the average payoffs after a deviation be not greater than the equilibrium payoff.

The following theorem characterizes the set of equilibrium payoffs of the infinitely repeated game.

Theorem 13.17 (The Folk Theorem for the infinitely repeated game Γ_∞) *The set of equilibrium payoffs of* Γ_∞ *is the set* $F \cap V$.

That is, every payoff vector that is in the convex hull of the payoffs $\{u(s), s \in S\}$ and is individually rational (i.e., that is not less than the minmax value $\overline{v}_i$ for each player i, which a player cannot be prevented from attaining) is an equilibrium payoff.

Note that the Folk Theorem for Γ_∞ and the Folk Theorem for Γ_T differ in two respects. First, in Γ_∞ we need not approximate a payoff to within ε; exact payoffs can be obtained. Second, in the finite repeated game, we required that for every player i there exist an equilibrium of the base game that gives player i a payoff that is strictly greater than his minmax value; this requirement is not needed for the Folk Theorem for Γ_∞. The differences between these two theorems are illustrated in the repeated Prisoner's Dilemma. In that example, for every $T \in \mathbb{N}$ the only equilibrium payoff of Γ_T is (1, 1) (see Example 13.1 on page 521), while according to Theorem 13.17, the set of equilibrium payoffs of Γ_∞ is the set W, shown in Figure 13.10.

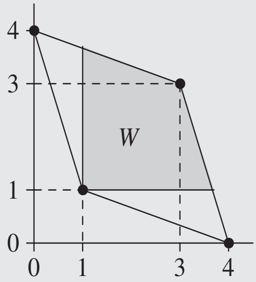

Figure 13.10 The set of equilibrium payoffs of the infinitely repeated Prisoner's Dilemma

Example 13.1 (*Continued*) Consider again the repeated Prisoner's Dilemma of Figure 13.9. We will show that (1, 2), for example, is an equilibrium payoff of the infinitely repeated game. We do so by constructing an equilibrium leading to this payoff. Note first that

$$(1,2) = \tfrac{5}{8}(1,1) + \tfrac{1}{8}(3,3) + \tfrac{2}{8}(0,4). \tag{13.40}$$

Define the pair of strategies $\tau^* = (\tau^*_{\mathrm{I}}, \tau^*_{\mathrm{II}})$ that repeatedly cycle through the action vectors

$$(D,D), (D,D), (D,D), (D,D), (D,D), (C,C), (C,D), (C,D), \tag{13.41}$$

unless a player deviates, in which case the other player switches to the punishment action D. Formally:

- Player I repeatedly cycles through the actions D, D, D, D, D, C, C, C.
- Player II repeatedly cycles through the actions D, D, D, D, D, C, C, D.
- If one of the players deviates, and fails to play the action he is supposed to play under this plan, the other player chooses D in every subsequent stage of the game, forever.

Direct inspection shows that τ^* is an equilibrium of the infinitely repeated game. By Equation (13.40) the average payoff at this equilibrium converges to (1, 2).

We can similarly obtain every payoff vector in $F \cap V$ that is representable as a weighted average of the vectors $(u(s))_{s\in S}$, with rational weights, as an equilibrium payoff of the infinitely repeated game.

When a payoff vector $x \in F \cap V$ cannot be represented as a weighted average with rational weights, we can use Theorem 13.11 to approximate x by way of a weighted average with rational weights. In other words, for every $k \in \mathbb{N}$ we can find rational coefficients $(\lambda_s(k))_{s\in S}$ with denominator k such that

$$\left| x - \sum_{s\in S} \lambda_s(k) u(s) \right| < \frac{1}{k}. \tag{13.42}$$

In this case, in the basic action plan, the players play in blocks, where the k-th block has k stages: in the first block the players play to obtain the average $\sum_{s\in S} \lambda_s(1)u(s)$, in the second block they play to obtain the average $\sum_{s\in S} \lambda_s(2)u(s)$, and so on. Recall that for every sequence $(z^k)_{k\in\mathbb{N}}$ of real numbers converging in the limit to z, the sequence of averages $(\frac{1}{n}\sum_{k=1}^{n} z^k)_{n\in\mathbb{N}}$ also converges to z.

Since the average payoff of the k-th block approaches x as k increases, the average payoff of the infinitely repeated game approaches x.

If one of the players deviates, the other player plays D in every stage from that stage on, forever, and hence no player can profit by deviating. ◀

The construction of the equilibrium strategy in the above example can be generalized to any repeated game, thus proving Theorem 13.17. The proof is left to the reader (Exercise 13.23).

As stated at the beginning of this section, one reason to study infinitely repeated games is to obtain insights into very long finitely repeated games. To present the connection between infinitely repeated games and finite games, we define the concept of ε-equilibrium of finite games.

Definition 13.18 *Let $\varepsilon > 0$, and let $T \in \mathbb{N}$. A strategy vector τ^* is an ε-*equilibrium *of Γ_T if for each player $i \in \mathbb{N}$ and any strategy $\tau_i \in \mathcal{B}_i^T$,*

$$\gamma_i^T(\tau^*) \geq \gamma_i^T(\tau_i, \tau_{-i}^*) - \varepsilon. \tag{13.43}$$

If τ^* is an ε-equilibrium of Γ_T, it is perhaps possible for a player to profit by deviating, but his profit will be no greater than ε. The smaller ε is, the less motivation a player has to deviate. When $\varepsilon = 0$, we recapitulate the definition of equilibrium, in which case no player has any motivation to deviate. If there is a cost for deviating, and the cost exceeds ε, then even at an ε-equilibrium, deviating is unprofitable. In this sense, ε-equilibria satisfy the property of being "almost stable," where "almost" is measured by ε.

For every strategy vector τ in Γ_∞, and every $T \in \mathbb{N}$, we can define the restriction of τ to the first T stages of the game. To avoid a plethora of symbols, we will denote such a restricted strategy vector by the same symbol, τ.

A stronger formulation of the Folk Theorem for Γ_∞ relates the equilibria of the infinitely repeated game to ε-equilibria in long finitely repeated games. The proof of the theorem is left to the reader (Exercise 13.25).

Theorem 13.19 *For every $\varepsilon > 0$ and every vector $x \in F \cap V$ there exist a strategy vector τ in Γ_∞ and $T_0 \in \mathbb{N}$ that satisfy the following:*

1. *τ is an equilibrium of Γ_∞.*
2. *τ is an ε-equilibrium of Γ_T, for all $T \geq T_0$.*

Example 13.1 (*Continued*) On page 522, we saw that the only equilibrium payoff of the T-stage repeated Prisoner's Dilemma is (1, 1). It follows that for every payoff vector $x \in F \cap V$ that is not (1, 1), the corresponding strategy vector τ constructed on page 540, which is an equilibrium of the infinitely repeated game, is not an equilibrium of the T-stage repeated game, for any $T \in \mathbb{N}$. However, for every $\varepsilon > 0$, for T sufficiently large, the strategy vector τ is an ε-equilibrium of the T-stage repeated game with an average payoff close to x. In other words, every payoff vector $x \in F \cap V$ can be supported by an ε-equilibrium in the T-stage repeated game, provided T is large enough (Exercise 13.24). ◀

As the following example shows, it is not the case that every equilibrium of Γ_∞ is an ε-equilibrium of every sufficiently long finitely repeated game.

Example 13.20 Consider the two-player zero-sum game in Figure 13.11.

Player II

Player I	L	R
T	0	−1
B	1	0

Figure 13.11 The base game in Example 13.20

The pure action B strictly dominates T, and R strictly dominates L. Elimination of strictly dominated actions reveals that the value of the (finitely or infinitely) repeated game is 0.

Consider the following pair of strategies $\tau^* = (\tau^*_{\mathrm{I}}, \tau^*_{\mathrm{II}})$ in the infinitely repeated game:

- τ^*_{I} instructs Player I to play T up to stage $(t_{\mathrm{II}})^2$ and to play B thereafter, where t_{II} is the first stage in which Player II plays R ($t_{\mathrm{II}} = \infty$ if Player II plays L in every stage).
- τ^*_{II} instructs Player II to play L up to stage $(t_{\mathrm{I}})^2$ and to play R thereafter, where t_{I} is the first stage in which Player I plays B ($t_{\mathrm{I}} = \infty$ if Player II plays T at every stage).

Thus, Player I plays T and checks whether Player II plays L (in which case the payoff is 0). As long as Player II plays L, Player I plays T. If at a certain stage (which we denote by t_{II}) Player II first plays R, Player I continues to play T for several stages (up to stage $(t_{\mathrm{II}})^2$) and from that stage on he punishes Player II by playing B in every subsequent stage. Player II's strategy is defined similarly, all things being equal.

At strategy vector τ^*, the players play (T, L) in every stage, and the payoff, in the infinitely repeated game, is 0. If one of the players (say Player I) deviates he may receive the higher payoff of 1 for several stages (the stages between t_{I} and $(t_{\mathrm{I}})^2$), but afterwards he can receive at most 0 in every stage. The upper limit of the average payoff of Player I in Γ_∞ is therefore less than or equal to 0 even when he deviates: he cannot profit by deviating. In particular, $(\tau^*_{\mathrm{I}}, \tau^*_{\mathrm{II}})$ is an equilibrium of the infinitely repeated game.

However, in finite games, $(\tau^*_{\mathrm{I}}, \tau^*_{\mathrm{II}})$ is not an ε-equilibrium for ε close to 0: for example, in the 99-stage repeated game, if Player I deviates from τ^*_{I}, and plays B from stage 10 onwards, his average payoff is $\frac{90}{99}$. It follows that a deviation yields Player I a profit of $\frac{90}{99}$. In general, in the T-stage game, playing against strategy τ^*_{II}, Player I has a deviation yielding him a payoff of $\frac{T-\lceil\sqrt{T}\rceil}{T}$ (Exercise 13.26). Similarly, playing against strategy τ^*_{I}, Player II has a deviation yielding him a payoff of $-\frac{T-\lceil\sqrt{T}\rceil}{T}$. It follows that when ε is sufficiently small, $(\tau^*_{\mathrm{I}}, \tau^*_{\mathrm{II}})$ is not an ε-equilibrium in any finite repeated game Γ_T, for large T. ◀

13.6 The discounted game

In the definition of the T-stage game, we assumed that every player seeks to maximize his average expected payoff at every stage of the game, or, equivalently, that every player seeks to maximize the expected sum of his payoffs. This means that if, say, John receives

\$10,000 today, and Paul receives \$10,000 in a year from now, their situations are considered identical. Such an assumption is not appealing: in reality, if John invests his \$10,000 in a bank account yielding, say, 5% annual interest, he will have \$10,500 in one year, and thus will be better off than Paul. This is the reason that economic models usually assume that players maximize not the sum of their payoffs over time, but the discounted sum of their payoffs, where the discount rate takes into account the interest that players can receive over time for their money.

Discounted repeated games are presented in this section. For mathematical convenience, we will consider only infinitely repeated games. The assumption that all games are infinite is applicable in realistic models, because when payoffs are time-discounted, payoffs in far-off stages have a negligible effect on the discounted sum of the payoffs. We will study the equilibria of such games, and compare them to the equilibria of the finitely and infinitely repeated games presented in the previous sections.

Definition 13.21 *Let* $\lambda \in [0, 1)$ *be a real number, and let* $\tau = (\tau_i)_{i \in N}$ *be a strategy vector in an infinitely repeated game. The* λ-discounted payoff *to player* i *under strategy vector* τ *is*

$$\gamma_i^\lambda(\tau) := \mathbf{E}_\tau \left[(1-\lambda) \sum_{t=1}^{\infty} \lambda^{t-1} u_i^t \right]. \tag{13.44}$$

The constant λ *is called the* discount factor.

The coefficient λ^{t-1} that multiplies the stage payoff u_i^t in Equation (13.44) expresses the fact that a payoff of \$1 tomorrow is equivalent to a payoff of \$$\lambda$ today, a payoff of \$1 in two days is equivalent to a payoff of \$$\lambda^2$ today, and so on. Since $(1-\lambda)\sum_{t=1}^{\infty} \lambda^{t-1} = 1$, the discounted payoff is the weighted average of the daily payoffs, where the weights decrease exponentially. When $\lambda = 0$, players' payoffs in Γ_λ equal their payoffs in the first stage of the game, and the discounted repeated game is essentially equivalent to the (one-stage) base game. When λ is close to zero, $1-\lambda$ (the weight associated with the first stage) is large relative to λ (the total weight associated with the payoffs in the subsequent stages), and the first-stage payoff is the most important one: players attach more importance to today's payoff, and are willing to forgo high payoffs in the future. When λ is close to 1, the weight associated with stage t is very close to that of stage $t+1$, and hence the players exhibit "patience": each player evaluates tomorrow's payoff almost as much as he evaluates today's payoff.

Since the sum total of the weights is 1, the λ-discounted payoff γ_i^λ may be viewed as an "expected payoff per stage." This can be seen in two different ways:

1. If the payoffs are 1 in each stage, we want the "average payoff" per stage to be 1, and indeed the discounted sum in this case is $(1-\lambda)\sum_{t=1}^{\infty} \lambda^{t-1} = 1$.
2. We can also interpret the discount factor λ as the probability that the game will continue to the next stage. In other words, at every stage there is probability $1-\lambda$ that the game will end, and probability λ that the game will continue. It follows that $\frac{1}{1-\lambda}$ is the expected number of stages to the end of the game, and the probability that the game will get to stage t is λ^{t-1}. With this interpretation, the sum on the right-hand side

of Equation (13.44) is the total expected payoff in the game divided by the expected number of stages.

Finally, since the definition in Equation (13.44) captures the per-stage payoff to player i, it allows us to compare equilibrium payoffs in discounted games, for different discount factors, and to compare these payoffs with equilibrium payoffs in finitely and infinitely repeated games.

Definition 13.22 *Let* $\Gamma = (N, (S_i)_{i \in N}, (u_i)_{i \in N})$ *be a base game, and let* $\lambda \in [0, 1)$. *The* discounted game Γ_λ *(with discount factor* λ*) corresponding to* Γ *is the game* $\Gamma_\lambda = (N, (\mathcal{B}_i^\infty)_{i \in N}, (\gamma_i^\lambda)_{i \in N})$.

It follows that a strategy vector τ^* is an equilibrium of Γ_λ if for each player $i \in \mathbb{N}$ and each strategy τ_i,

$$\gamma_i^\lambda(\tau^*) \geq \gamma_i^\lambda(\tau_i, \tau_{-i}^*). \tag{13.45}$$

In this case, the vector $\gamma_i^\lambda(\tau^*)$ is an *equilibrium payoff* of Γ_λ. The minmax value of each player i in Γ_λ is his minmax value in the base game Γ (Exercise 13.31), and an equilibrium payoff of a player is at least his minmax value (Theorem 5.42 on page 180). Therefore, $\gamma_i^\lambda(\tau^*) \geq \overline{v}_i$ for each player $i \in N$.

So far we have seen two ways to model long repeated games, using the infinitely repeated game Γ_∞ and using finite repeated games with duration T that increases to infinity. As we have seen in point 2 above, in a λ-discounted model we can interpret the quantity $\frac{1}{1-\lambda}$ as the expected duration of the game. Since this quantity goes to infinity as λ goes to 1, a third way to model a long repeated game is by λ-discounted games with a discount factor λ that goes to 1. A natural question that arises concerns the limit set of the set of λ-discounted equilibrium payoffs as λ goes to 1. In view of the Folk Theorem for infinitely repeated games, can we prove an analog result for discounted games, that is, is it true that every vector $x \in F \cap V$ is the limit of λ-discounted equilibrium payoffs, as the discount factor goes to 1? The following example shows that this is not the case.

Example 13.23 Consider the three-player base game given in Figure 13.12, in which Player I has three actions $\{T, M, B\}$, Player II has two actions $\{L, R\}$, and Player III is a dummy player who has only one action, which has no effect on the payoffs (and is not mentioned throughout the example).

		Player II	
		L	R
Player I	T	0, 2, 5	0, 0, 0
	M	0, 1, 0	2, 0, 5
	B	1, 1, 0	1, 1, 0

Figure 13.12 The base game in Example 13.23

The minmax values of the players are 1, 1, and 0 respectively, and the set $F \cap V$ of the feasible and individually rational payoffs is the line segment $[(1, 1, 0) - (1, 1, 5)]$ (verify!). We will now show that $(1, 1, 0)$ is the only equilibrium payoff in the discounted game Γ_λ, for any discount factor $\lambda \in [0, 1)$.

Let τ^* be an equilibrium of the discounted game Γ_λ. We first show that $\gamma^\lambda_{\mathrm{I}}(\tau^*) = \gamma^\lambda_{\mathrm{II}}(\tau^*) = 1$. Indeed, $\gamma^\lambda_i(\tau^*) \geq \overline{v}_i = 1$ for $i \in \{\mathrm{I}, \mathrm{II}\}$. On the other hand, the sum of the payoffs to Players I and II is at most 2 in all entries of the payoff matrix, and therefore $\gamma^\lambda_{\mathrm{I}}(\tau^*) + \gamma^\lambda_{\mathrm{II}}(\tau^*) \leq 2$. Consequently, $\gamma^\lambda_{\mathrm{I}}(\tau^*) = \gamma^\lambda_{\mathrm{II}}(\tau^*) = 1$, as claimed.

Since the sum of payoffs for Players I and II at (M, L) and (T, R) is strictly less than 2, these two pairs of actions are chosen with probability 0 at the equilibrium τ^*; otherwise, we would have $\gamma^\lambda_{\mathrm{I}}(\tau^*) + \gamma^\lambda_{\mathrm{II}}(\tau^*) < 2$, which contradicts $\gamma^\lambda_{\mathrm{I}}(\tau^*) = \gamma^\lambda_{\mathrm{II}}(\tau^*) = 1$.

For any $t \geq 0$ and any history $h^t \in H(t)$ denote by $\gamma^\lambda_i(\tau^* \mid h^t)$ the conditional discounted future payoff of player i (from stage $t + 1$ on) given the history h^t, under the equilibrium τ^*:

$$\gamma^\lambda_i(\tau^* \mid h^t) := \mathbf{E}_{\tau^*}\left[(1 - \lambda) \sum_{j=1}^{\infty} \lambda^{j-1} u_i^{t+j} \mid h^t \right]. \tag{13.46}$$

The arguments provided above show that $\gamma^\lambda_i(\tau^* \mid h^t) = 1$ for $i \in \{\mathrm{I}, \mathrm{II}\}$, for any history h^t that has positive probability under τ^*. To prove that $\gamma^\lambda_{\mathrm{III}}(\tau^*) = 0$ we will show that the pairs of actions (T, L) and (M, R) are chosen under τ^* with probability 0. Assume by contradiction that the action pair (T, L) is chosen with positive probability $\alpha > 0$ at some stage $t \geq 0$ after the history $h^t \in H(t)$. Since (M, L) and (T, R) are played with probability 0, it follows that at the history h^t, the action pair (B, L) is played with probability $1 - \alpha$. Therefore,

$$\begin{aligned}
1 &= \gamma^\lambda_{\mathrm{II}}(\tau^* \mid h^t) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (13.47)\\
&= \alpha \left((1 - \lambda) \times 2 + \lambda \times \gamma^\lambda_{\mathrm{II}}(\tau^* \mid (h^t, (T, L)))\right)\\
&\quad + (1 - \alpha)\left((1 - \lambda) \times 1 + \lambda \times \gamma^\lambda_{\mathrm{II}}(\tau^* \mid (h^t, (B, L)))\right) \qquad\qquad\qquad\qquad (13.48)\\
&= \alpha \left((1 - \lambda) \times 2 + \lambda \times 1\right) + (1 - \alpha)\left((1 - \lambda) \times 1 + \lambda \times 1\right), \qquad\qquad (13.49)
\end{aligned}$$

which implies that $\alpha = 0$, in contradiction to our assumption that $\alpha > 0$. This proves that the action pair (T, L) is played with probability 0 under τ^*, and similarly the action pair (M, L) is played with probability 0 under τ^*. This concludes the proof that $\gamma^\lambda(\tau^*) = (1, 1, 0)$.

The fact that at the only equilibrium payoff every player's payoff is his minmax value is a coincidence. Indeed, replacing the payoffs 5 in Figure 13.12 by (-5) does not affect our proof that $(1, 1, 0)$ is the only equilibrium payoff, while the minmax value of Player III changes to (-5). ◀

The Folk Theorem for finitely repeated games required a technical condition on the base game: for every player i there exists an equilibrium $\beta(i)$ of the base game that yields player i a payoff higher than his minmax value. In our construction of equilibria in the repeated game, the mixed actions $(\beta(i))_{i=1}^n$ were played in the last stages of the game, to ensure that a player who deviates along the play will lose when punished. To obtain a Folk Theorem for discounted games one also needs a technical condition on the base game for the same purpose. The condition that we will require is weaker than the one that appears in the Folk Theorem for finite games (Exercise 13.41).

Theorem 13.24 (The Folk Theorem for discounted games) *Let Γ be a base game in which there exists a vector $\widehat{x} \in F \cap V$ that satisfies $\widehat{x}_i > \overline{v}_i$ for every player $i \in N$. For every $\varepsilon > 0$ there exists $\lambda_0 \in [0, 1)$ such that for every $\lambda \in [\lambda_0, 1)$ and every vector*

$x \in F \cap V$, there exists an equilibrium τ^* of Γ_λ satisfying[6]

$$\|\gamma^\lambda(\tau^*) - x\|_\infty < \varepsilon. \tag{13.50}$$

The condition that appears in the statement of the theorem ensures that there is a convex combination of the entries of the payoff matrix, using rational weights, that is close to x, and yields each player a payoff that is strictly higher than his minmax value. If, as we did in the proof of Theorem 13.17, we construct a strategy vector with a basic plan in which the players play according to this convex combination, then, for λ sufficiently close to 1, for every t, the λ-discounted payoff of every player, from stage t on, is strictly higher than his minmax value, allowing the players to punish a deviator. The reader is asked to complete the details of the proof (Exercise 13.42).

13.7 Uniform equilibrium

As we said before, players may not always know the number of stages a repeated game will have:

- A young professional baseball player knows that at a certain age he will retire from the sport, but does not know exactly when that day will come.
- We are all players in "the game of life," whose length is unknown and differs among the players.

Similarly, players do not always know the discount factor of the game:

- Although the prime interest rate is common knowledge, we do not know what the interest rate will be next year, in two years, or a decade from today.
- Suppose the government is interested in selling a state-owned company. What discount rate should be used? Computing a reasonable discount rate in such cases can be very complicated.

In the examples above, the discount rate and the exact length of the game are unknown. How should a player play in this case? In this section we will present concepts enabling us to study this question, and to arrive at results that are independent of the exact value of the discount factor, or the exact length of the game. To do so, we introduce the concept of *uniform equilibrium*, first for discounted games, then for finite games, and later we will see the relation between the two.

Definition 13.25 *A strategy vector τ^* is called a* uniform equilibrium for discounted games *if* $\lim_{\lambda \to 1} \gamma^\lambda(\tau^*)$ *exists, and there exists $\lambda_0 \in [0, 1)$ such that τ^* is an equilibrium of Γ_λ for every discount factor $\lambda \in [\lambda_0, 1)$. The limit* $\lim_{\lambda \to 1} \gamma^\lambda(\tau^*)$ *is called a* uniform equilibrium payoff for discounted games.

τ^* is therefore a uniform equilibrium for discounted games if it is an equilibrium of every game in which the discount factor is sufficiently close to 1; that is, the players are sufficiently patient.

---

[6] Recall that the maximum norm over $\mathbb{R}^n$ is defined as $\|x\|_\infty = \max_{i=1,2,\ldots,n} |x_i|$ for every vector $x \in \mathbb{R}^n$.

Do uniform equilibria for discounted games exist? As we will see (Theorem 13.27), there are many uniform equilibria for discounted games.

Example 13.26 Consider the two-player game in Figure 13.13.

Player II

Player I		*L*	*R*
	T	3, 1	0, 0
	B	1, 2	4, 3

Figure 13.13 The payoff matrix of the game in Example 13.26

The minmax value of Player I is $\overline{v}_{\mathrm{I}} = 2$, and the punishment strategy against him is $[\frac{2}{3}(L), \frac{1}{3}(R)]$. Player II's minmax value is $\overline{v}_{\mathrm{II}} = 1$, and the punishment strategy against him is T. We will show that in the example above,

$$\left(3\tfrac{1}{2}, 2\tfrac{5}{8}\right) = \tfrac{7}{8} \times (4, 3) + \tfrac{1}{8} \times (0, 0) \tag{13.51}$$

is a uniform equilibrium payoff for discounted games. To do so, we will show that the following pair of strategies is a discounted equilibrium for every discount factor sufficiently close to 1:

- Player I plays B at the first stage, and as long as Player II plays R, Player I repeatedly cycles through the following sequence of actions: B, B, B, B, B, B, B, T (the action B played at the first stage is the beginning of the first cycle).
- Player II plays R at the first stage, and as long as Player I cycles through the sequence of actions B, B, B, B, B, B, B, T, Player II plays R.

If neither player deviates from this strategy, the discounted sum of the payoffs of the first eight stages of the game is

$$(1 + \lambda + \lambda^2 + \cdots + \lambda^6)(4, 3) + \lambda^7(0, 0) = \frac{1 - \lambda^7}{1 - \lambda} \cdot (4, 3). \tag{13.52}$$

Therefore, the discounted payoff is

$$\begin{aligned}
&(1 - \lambda)((4, 3) + \lambda(4, 3) + \lambda^2(4, 3) + \cdots + \lambda^6(4, 3) + \lambda^7(0, 0) + \lambda^8(4, 3) \\
&\quad + \lambda^9(4, 3) + \cdots + \lambda^{14}(4, 3) + \lambda^{15}(0, 0) + \cdots) \\
&= (1 - \lambda) \times \frac{1 - \lambda^7}{1 - \lambda}(1 + \lambda^8 + \lambda^{16} + \cdots) \cdot (4, 3) && (13.53) \\
&= \frac{1 - \lambda^7}{1 - \lambda^8} \cdot (4, 3). && (13.54)
\end{aligned}$$

Applying L'Hôpital's Rule, the limit of this value, as λ approaches 1, is

$$\lim_{\lambda \to 1} \left(\frac{1 - \lambda^7}{1 - \lambda^8} \cdot (4, 3) \right) = \left(\lim_{\lambda \to 1} \frac{-7\lambda^6}{-8\lambda^7} \right) \times (4, 3) = \tfrac{7}{8} \times (4, 3) = \left(3\tfrac{1}{2}, 2\tfrac{5}{8}\right). \tag{13.55}$$

Neither player can profit by deviating in the stages in which the players play (B, R), because in these stages each receives his maximal possible payoffs. To guarantee that neither player can profit by deviating in the stages in which the players play (T, R), we add the following punishments:

- If Player II deviates for the first time in stage t, from stage $t+1$ onwards Player I always plays T (which is Player I's punishment strategy against Player II).
- If Player I deviates for the first time in stage t, from stage $t+1$ onwards Player II always plays the mixed action $[\frac{2}{3}(L), \frac{1}{2}(R)]$ (which is Player II's punishment strategy against Player I).

We next seek discount factors λ for which this strategy vector is a λ-discounted equilibrium. If Player I deviates in stage t, where the players are supposed to play (T, R), he receives in that stage a payoff of 4 instead of 0, for a net profit of 4. In contrast, from the next stage onwards, his expected payoff is 2. Player I's λ-discounted payoff from stage t onwards is therefore[7]

$$(1-\lambda)(4+\lambda\times 2+\lambda^2\times 2+\lambda^3\times 2+\cdots)=4(1-\lambda)+2\lambda=4-2\lambda. \tag{13.56}$$

If Player I had not deviated in stage t, his discounted payoff from stage t onwards would be

$$(1-\lambda)\times 0+\lambda\times\frac{1-\lambda^7}{1-\lambda^8}\times 4, \tag{13.57}$$

because if in stage t the players are supposed to play (T, R), then a new cycle of length 8 begins in stage $t+1$, and therefore the λ-discounted payoff from stage $t+1$ onwards equals (up to multiplication by λ^t) the λ-discounted payoff from the first stage. The deviation is unprofitable only if the payoff, when no deviation occurs, is greater than or equal to the payoff when a deviation occurs:

$$\lambda\times\frac{1-\lambda^7}{1-\lambda^8}\times 4\geq 4-2\lambda. \tag{13.58}$$

Multiplying both sides of the expression by $1-\lambda^8$, we deduce that the following must hold:

$$4\lambda-4\lambda^8\geq 4-4\lambda^8-2\lambda+2\lambda^9. \tag{13.59}$$

For $\lambda = 1$, both sides of the expression equal zero. Differentiating the left-hand side and setting $\lambda = 1$ yields $4-32=-28$, while differentiating the right-hand side and setting $\lambda = 1$ yields $-32-2+18=-16$. Therefore, an interval $(\lambda_0, 1)$ exists such that for every discount factor λ in the interval, the left-hand side of Equation (13.59) is greater than the right-hand side of Equation (13.59). One can check that the inequality in Equation (13.59) holds as a strict inequality for every $\lambda \in (0.615, 1)$.

If Player II deviates in stage t, where the players are supposed to play (T, R), he receives a payoff of 1 instead of 0 in that stage, for a net profit of 1. In contrast, from that stage onwards his payoff is bounded by 1. It follows that the λ-discounted payoff of Player II from stage t onwards is at most 1. In contrast, if Player II does not deviate, his payoff from stage t onwards is $(1-\lambda)\times 0+\lambda\times\frac{1-\lambda^7}{1-\lambda^8}\times 3$. Deviating is not profitable if

$$(1-\lambda)\times 0+\lambda\times\frac{1-\lambda^7}{1-\lambda^8}\times 3\geq 1. \tag{13.60}$$

It can be shown that this holds for all $\lambda \geq 0.334$. We deduce from this that the pair of strategies defined above form a λ-discounted equilibrium for all $\lambda > \max\{0.334, 0.615\} = 0.615$; hence it is a uniform equilibrium for discounted games. ◀

---

[7] For all $a, b \in \mathbb{R}$, $(1-\lambda)\left(\lambda^{t-1}a+\lambda^t b+\lambda^{t+1}b+\cdots\right)=\lambda^{t-1}((1-\lambda)a+\lambda b)$.

Theorem 13.27 (The Folk Theorem for uniform equilibrium in discounted games) *Let* Γ *be a base game in which there exists a vector* $\widehat{x} \in F \cap V$ *that satisfies* $\widehat{x}_i > \overline{v}_i$ *for every player* $i \in N$*. For every* $\varepsilon > 0$*, and every* $x \in F \cap V$*, there exists a strategy vector* τ^* *in the discounted repeated game such that:*

1. τ^* *is a uniform equilibrium for discounted games.*
2. $\| \lim_{\lambda \to 1} \gamma^{\lambda}(\tau^*) - x \|_{\infty} < \varepsilon$.

In words, for each $x \in F \cap V$ there exists a uniform equilibrium for discounted games τ^* satisfying the property that the limit of discounted payoffs $\lim_{\lambda \to 1} \gamma^{\lambda}(\tau^*)$ is approximately x. The strategy vector τ^* satisfying the conditions of the theorem, similarly to the case in the proof of Theorem 13.17 (page 539), is of the grim-trigger type, with a basic plan that ensures that the payoff vector is close to x. The complete proof of this theorem is left to the reader (Exercise 13.44).

The concept of uniform equilibrium can also be defined for long finite games. We will see in Theorem 13.32 that the two concepts of uniform equilibrium are related.

Definition 13.28 *Let* $\varepsilon \geq 0$*. A strategy vector* τ^* *in an infinitely repeated game is a* uniform ε-equilibrium for finite games *if the limit* $\lim_{T \to \infty} \gamma^T(\tau^*)$ *exists, and there exists* $T_0 \in \mathbb{N}$ *such that* τ^* *is an* ε*-equilibrium of* Γ_T*, for every* $T \geq T_0$*. The limit* $\lim_{T \to \infty} \gamma^T(\tau^*)$ *is called a* uniform ε-equilibrium payoff.

At every uniform 0-equilibrium for finite games (i.e., the case in which $\varepsilon = 0$), from some stage onwards, in every stage, the players play an equilibrium of the base game (Exercise 13.48). Consequently, the set of uniform 0-equilibrium payoffs is the convex hull of the set of Nash equilibrium payoffs of the base game. For $\varepsilon > 0$, however, the set of uniform ε-equilibrium payoffs is much larger; that is, the Folk Theorem holds.

Theorem 13.29 (The Folk Theorem for uniform equilibrium in finite games) *For every* $\varepsilon > 0$*, and every* $x \in F \cap V$*, there exists a strategy vector* τ^* *such that:*

1. τ^* *is a uniform* ε*-equilibrium for finite games:*
2. $\| \lim_{T \to \infty} \gamma^T(\tau^*) - x \|_{\infty} < \varepsilon$.

The proof of the theorem, which is similar to the proof of Theorem 13.17 (page 539), is left to the reader (Exercise 13.49).

We now turn our attention to comparing the concepts of uniform equilibrium for discounted games and uniform ε-equilibrium for finite games. For this purpose, we will first find a connection between finite averages and discounted sums.

Theorem 13.30 *Let* $(x_t)_{t=1}^{\infty}$ *be a bounded sequence of numbers. Denote the average of the first* T *elements of this sequence by*

$$S_T = \frac{1}{T} \sum_{t=1}^{T} x_t, \qquad \forall T \in \mathbb{N}, \tag{13.61}$$

and the discounted sum by

$$A(\lambda) = (1-\lambda)\sum_{t=1}^{\infty}\lambda^{t-1}x_t, \qquad \forall\lambda \in [0,1). \tag{13.62}$$

Also denote

$$\alpha_T(\lambda) = (1-\lambda)^2\lambda^{T-1}T, \qquad \forall T \in \mathbb{N}, \forall\lambda \in [0,1). \tag{13.63}$$

Then, for all $\lambda \in [0, 1)$,

$$A(\lambda) = \sum_{T=1}^{\infty}\alpha_T(\lambda)S_T. \tag{13.64}$$

Note that $\sum_{T=1}^{\infty}\alpha_T(\lambda) = 1$:

$$\sum_{T=1}^{\infty}\alpha_T(\lambda) = (1-\lambda)^2\sum_{T=1}^{\infty}T\lambda^{T-1} \tag{13.65}$$

$$= (1-\lambda)^2\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\sum_{T=1}^{\infty}\lambda^T\right) \tag{13.66}$$

$$= (1-\lambda)^2\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{1}{1-\lambda}\right) = 1, \tag{13.67}$$

where Equation (13.66) follows from the Bounded Convergence Theorem (see Theorem 16.4 in Billingsley [1999]). Thus, Equation (13.64) states that $A(\lambda)$ is a weighted average of $(S_T)_{T\in\mathbb{N}}$.

Proof: The proof of the theorem is accomplished by the following sequence of equalities:

$$A(\lambda) = (1-\lambda)\sum_{t=1}^{\infty}\lambda^{t-1}x_t \tag{13.68}$$

$$= (1-\lambda)\sum_{t=1}^{\infty}\left(\sum_{k=t}^{\infty}(\lambda^{k-1}-\lambda^k)\right)x_t \tag{13.69}$$

$$= (1-\lambda)\sum_{k=1}^{\infty}\left((\lambda^{k-1}-\lambda^k)\sum_{t=1}^{k}x_t\right) \tag{13.70}$$

$$= \sum_{k=1}^{\infty}\left((1-\lambda)(\lambda^{k-1}-\lambda^k)kS_k\right) \tag{13.71}$$

$$= \sum_{k=1}^{\infty}(1-\lambda)^2\lambda^{k-1}kS_k \tag{13.72}$$

$$= \sum_{k=1}^{\infty}\alpha_k(\lambda)S_k. \tag{13.73}$$

Equation (13.70) follows by changing the order of summation (why can the order of summation be changed in this case?), and Equation (13.70) follows from Equation (13.61). □

The next theorem is a consequence of Theorem 13.30.

Theorem 13.31 (Hardy and Littlewood) *Every bounded sequence of real numbers* $(x_t)_{t=1}^{\infty}$ *satisfies*

$$\liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x_t \le \liminf_{\lambda\to 1}\sum_{t=1}^{\infty}(1-\lambda)\lambda^{t-1}x_t \tag{13.74}$$

$$\le \limsup_{\lambda\to 1}\sum_{t=1}^{\infty}(1-\lambda)\lambda^{t-1}x_t \tag{13.75}$$

$$\le \limsup_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} x_t. \tag{13.76}$$

In particular, if the limit $\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} x_t$ *exists, then the limit* $\lim_{\lambda\to 1}\sum_{t=1}^{\infty}(1-\lambda)\lambda^{t-1}x_t$ *also exists, and both limits are equal.*

Using the notation of Theorem 13.30, Theorem 13.31 states that

$$\liminf_{T\to\infty} S_T \le \liminf_{\lambda\to 1} A(\lambda) \le \limsup_{\lambda\to 1} A(\lambda) \le \limsup_{T\to\infty} S_T. \tag{13.77}$$

Proof: We will prove Equation (13.76):

$$\limsup_{\lambda\to 1}\sum_{t=1}^{\infty}(1-\lambda)\lambda^{t-1}x_t \le \limsup_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} x_t. \tag{13.78}$$

The proof of Equation (13.74) can be accomplished in a similar manner, or by considering the sequence $(y_t)_{t\in\mathbb{N}}$ defined by $y_t := -x_t$ for every $t \in \mathbb{N}$. Equation (13.75) requires no proof. Theorem 13.30 implies that for all $T_0 \in \mathbb{N}$,

$$A(\lambda) = \sum_{T=1}^{\infty}\alpha_T(\lambda)S_T = \sum_{T=1}^{T_0-1}\alpha_T(\lambda)S_T + \sum_{T=T_0}^{\infty}\alpha_T(\lambda)S_T. \tag{13.79}$$

Denote $C := \limsup_{T\to\infty} S_T$, and let $\varepsilon > 0$ be any positive real number. Let T_0 be sufficiently large such that $S_T \le C+\varepsilon$ for all $T \ge T_0$. Note that

$$\sum_{T=1}^{T_0}\alpha_T(\lambda) = (1-\lambda)^2\sum_{T=1}^{T_0}\lambda^{T-1}T < (1-\lambda)^2(T_0)^2, \tag{13.80}$$

where the last inequality follows from the fact that $\lambda \in [0,1)$. In particular, when λ approaches 1 the sum $\sum_{T=1}^{T_0-1}\alpha_T(\lambda)$ approaches 0, and therefore the first sum in

Equation (13.79) also converges to 0. The second sum is bounded by $C + \varepsilon$. We therefore have

$$\limsup_{\lambda\to 1} A(\lambda) \leq C + \varepsilon = \limsup_{T\to\infty} S_T + \varepsilon. \tag{13.81}$$

Since this inequality holds for all $\varepsilon > 0$, we deduce that $\limsup_{\lambda\to 1} A(\lambda) \leq \limsup_{T\to\infty} S_T$, which is what we wanted to prove. □

Analogously to the definition for finite games (Definition 13.18), a strategy vector τ^* is an ε-equilibrium in the discounted game Γ_λ if no player can profit more than ε by deviating:

$$\gamma_i^\lambda(\tau^*) \geq \gamma_i^\lambda(\tau_i, \tau_{-i}^*) - \varepsilon, \quad \forall i \in N, \forall \tau_i \in \mathcal{B}_i^\infty. \tag{13.82}$$

Theorem 13.31 enables us to establish the following connection between uniform ε-equilibria for finite games and uniform ε-equilibria for discounted games.

Theorem 13.32 *Let τ^* be a uniform ε-equilibrium for finite games. Then, for every $\delta > 0$, there exists $\lambda_0 \in [0, 1)$ such that for every $\lambda \in [\lambda_0, 1)$, the strategy vector τ is an $(\varepsilon + 2\delta)$-equilibrium of the discounted game with discount factor λ: for every player $i \in \mathbb{N}$, and every strategy τ_i,*

$$\gamma_i^\lambda(\tau^*) \geq \gamma_i^\lambda(\tau_i, \tau_{-i}^*) - \varepsilon - 2\delta. \tag{13.83}$$

Proof: Let τ^* be a uniform ε-equilibrium for finite games, and let $\delta > 0$. Recall that M is a bound on the payoffs of the base game, and denote by $C := \lim_{T\to\infty} \gamma_i^T(\tau^*)$ the limit of the payoffs in the finite games. By Theorem 13.31,

$$C = \lim_{\lambda\to 1} \gamma_i^\lambda(\tau^*). \tag{13.84}$$

Let T_0 be sufficiently large such that for each $T \geq T_0$, one has (a) the strategy vector τ^* is an ε-equilibrium of the T-stage game, and (b) $|C_i - \gamma_i^T(\tau^*)| < \delta$.

Let λ_0 be sufficiently close to 1 such that $\sum_{T=1}^{T_0} \alpha_T(\lambda_0) \leq \frac{\delta}{M}$ (see Equation (13.80)). Let i be a player, and let τ_i be a strategy of player i. We will show that for λ sufficiently close to 1, player i cannot profit more than $\varepsilon + \delta$ by deviating to any strategy τ_i. Denote the expected payoff in stage t, when player i deviates to τ_i, by

$$x_t = \mathbf{E}_{\tau_i, \tau_{-i}^*}[u_i(a_t)]. \tag{13.85}$$

The average of $x_1, x_2, \ldots, x_T$ equals the payoff under (τ_i, τ_{-i}^*) in the T-stage game:

$$\gamma_i^T(\tau_i, \tau_{-i}^*) = \frac{\sum_{t=1}^T x_t}{T}. \tag{13.86}$$

For each $\lambda \in [\lambda_0, 1)$,

$$\gamma_i^\lambda(\tau_i, \tau_{-i}^*) = \sum_{T=1}^{\infty} \alpha_T(\lambda)\gamma_i^T(\tau_i, \tau_{-i}^*) \tag{13.87}$$

$$= \sum_{T=1}^{T_0-1} \alpha_T(\lambda)\gamma_i^T(\tau_i, \tau_{-i}^*) + \sum_{T=T_0}^{\infty} \alpha_T(\lambda)\gamma_i^T(\tau_i, \tau_{-i}^*) \tag{13.88}$$

$$\leq \delta + \sum_{T=T_0}^{\infty} \alpha_T(\lambda)\gamma_i^T(\tau_i, \tau_{-i}^*) \tag{13.89}$$

$$\leq \delta + \sum_{T=T_0}^{\infty} \alpha_T(\lambda)\gamma_i^T(\tau^*) + \varepsilon \tag{13.90}$$

$$\leq \delta + C + \delta + \varepsilon \tag{13.91}$$

$$= \lim_{\lambda \to 1} \gamma_i^\lambda(\tau^*) + \varepsilon + 2\delta. \tag{13.92}$$

Equation (13.87) holds by Theorem 13.30, Equation (13.89) holds because $\lambda \in [\lambda_0, 1)$ and by the choice of λ_0, and Equation (13.90) holds because τ^* is an ε-equilibrium for every $T \geq T_0$. Equation (13.91) holds because $|C_i - \gamma_i^T(\tau^*)| < \delta$ for every $T \geq T_0$, and Equation (13.92) follows from Equation (13.84). It follows that τ^* is an $(\varepsilon + 2\delta)$-equilibrium of the λ-discounted game, and this holds for all $\lambda \in [\lambda_0, 1)$. □

We have already seen in Example 13.20 that an equilibrium of Γ_∞ is not necessarily an ε-equilibrium of long finite games, and therefore not necessarily a uniform ε-equilibrium for finite games. The following example shows that a uniform equilibrium for discounted games is not necessarily a uniform ε-equilibrium for finite games, or an equilibrium of Γ_∞.

Example 13.33 Let $(x_t)_{t=1}^\infty$ be a sequence of zeros and ones satisfying

$$\limsup_{T\to\infty} \frac{\sum_{t=1}^T x_t}{T} > \limsup_{\lambda\to 1} (1-\lambda)\sum_{t=1}^{\infty} \lambda^{t-1} x_t. \tag{13.93}$$

For details on how to construct such a sequence, see Exercise 13.50. Let c be a real number satisfying

$$\limsup_{T\to\infty} \frac{\sum_{t=1}^T x_t}{T} > c > \limsup_{\lambda\to 1} (1-\lambda)\sum_{t=1}^{\infty} \lambda^{t-1} x_t. \tag{13.94}$$

Consider the two-player game in Figure 13.14. In this game, the payoff to Player II is 2, under every action vector. As we will now show, $(c, 2)$ is a uniform equilibrium payoff of discounted repeated games, but is not a uniform ε-equilibrium payoff of finite games, for $\varepsilon > 0$ sufficiently small. Since under every circumstance Player II receives 2 in every stage of the repeated game, to prove that a pair of strategies is an equilibrium it is sufficient to show that Player I cannot profit by deviating.

		Player II		
		D	E	F
Player I	A	0, 2	0, 2	c, 2
	B	0, 2	1, 2	c, 2

Figure 13.14 The payoff matrix of the game in Example 13.33

Define the following strategy σ_{II} of Player II:

- In the first stage, play F.
- If in the first stage Player I played A, play F in all of the remaining stages of the game.
- If in the first stage Player I played B, play D or E in all of the remaining stages of the game, according to the above-mentioned sequence $(x_t)_{t=1}^{\infty}$: if $x_t = 0$, play D in stage t, and if $x_t = 1$, play E in stage t.

The strategy σ_{II} does not depend on Player I's actions after the first stage. For Player I, therefore, every strategy σ_{I} is weakly dominated by the strategy in which Player I's action in the first stage is the same as that of σ_{I}, and from the second stage onwards his action is always B. It follows that Player I's best reply to σ_{II} is either σ_{I}^{A}, where Player I plays A in the first stage, and B in every other stage, or σ_{I}^{B}, where he plays B in every stage, including the first stage.

The strategy vector $(\sigma_{\mathrm{I}}^{A}, \sigma_{\mathrm{II}})$ is a uniform equilibrium for discounted games with payoff $(c, 2)$. To see this, note that since $\gamma_{\mathrm{I}}^{\lambda}(\sigma_{\mathrm{I}}^{A}, \sigma_{\mathrm{II}}) = c$, while $\gamma_{\mathrm{I}}^{\lambda}(\sigma_{\mathrm{I}}^{B}, \sigma_{\mathrm{II}}) = (1-\lambda)\sum_{t=1}^{\infty} \lambda^{t-1} x_t$, Equation (13.94) implies that for a discount factor sufficiently close to 1, one has $\gamma_{\mathrm{I}}^{\lambda}(\sigma_{\mathrm{I}}^{A}, \sigma_{\mathrm{II}}) > \gamma_{\mathrm{I}}^{\lambda}(\sigma_{\mathrm{I}}^{B}, \sigma_{\mathrm{II}})$, and therefore Player I has no profitable deviation.

We next show that $(\sigma_{\mathrm{I}}^{A}, \sigma_{\mathrm{II}})$ is not a uniform ε-equilibrium for finite games, for $\varepsilon > 0$ sufficiently small. Set $\varepsilon_0 := \frac{1}{2}\left(\limsup_{T\to\infty} \frac{\sum_{t=1}^{T} x_t}{T} - c\right)$. We will show that there exists an increasing sequence $(T_k)_{k\in\mathbb{N}}$ such that $(\sigma_{\mathrm{I}}^{A}, \sigma_{\mathrm{II}})$ is not an ε-equilibrium of the T_k-stage game, for every $k \in \mathbb{N}$ and every $\varepsilon \in (0, \varepsilon_0)$. By Equation (13.94), there exists an increasing sequence $(T_k)_{k\in\mathbb{N}}$ such that for every $k \in \mathbb{N}$,

$$\gamma_{\mathrm{I}}^{T_k}\left(\sigma_{\mathrm{I}}^{B}, \sigma_{\mathrm{II}}\right) = \frac{\sum_{t=1}^{T_k} x_t}{T_k} > c + \varepsilon_0 > c = \gamma_{\mathrm{I}}^{T_k}\left(\sigma_{\mathrm{I}}^{A}, \sigma_{\mathrm{II}}\right). \tag{13.95}$$

Therefore, for every $k \in \mathbb{N}$, by deviating in the T_k-stage game to σ_{I}^{B}, Player I's profit is more than ε_0. It follows that $(\sigma_{\mathrm{I}}^{A}, \sigma_{\mathrm{II}})$ is not an ε-equilibrium in Γ_{T_k} for every $k \in \mathbb{N}$ and every $\varepsilon \in (0, \varepsilon_0]$. We further note that it follows from this discussion that $(\sigma_{\mathrm{I}}^{A}, \sigma_{\mathrm{II}})$ is also not an equilibrium in the infinitely repeated game (Exercise 13.52). ◀

13.8 Discussion

There is a wealth of literature on repeated games, and many variations of this model have been studied. One line of inquiry has focused on the subject of punishment. The equilibrium strategies we have defined in this section are unforgiving: once a player deviates, he is punished by the other player for the rest of the game. Because a punishment strategy is liable to lower the payoff of not only the player who is being punished but also other players in the game, it is reasonable to ask whether players whose interests are harmed by a punishment strategy will join in implementing it. Considerations such as these have led to the study of subgame perfect equilibria in repeated games (for a discussion on the notion of subgame perfect equilibrium in an extensive-form game, see Section 7.1 on page 252). In repeated games, a strategy vector τ^* is a subgame perfect equilibrium if after every finite history (whether or not the players arrive at that history if they implement τ^*), the play of the game that ensues from that stage onwards is an equilibrium of the subgame starting at that point. A proof of the Folk Theorem under this definition of equilibrium appears in Aumann and Shapley [1994], Rubinstein [1979], Fudenberg and Maskin [1986], and Gossner [1995].

There are several other variations on the theme of repeated games that have been studied in the literature. These include what happens when: (1) players do not observe the actions implemented by other players, and instead receive only a signal that depends on the actions of all the players (see, e.g., Lehrer [1989], [1990], and [1992], and Gossner and Tomala [2007]); (2) players do not know their payoff functions (see, e.g., Megiddo [1980]); and (3) at the start of the game, a payoff function is chosen from a set of possible payoff functions, and the players receive partial information regarding which payoff function is chosen (see Aumann and Maschler [1995] and Section 14.7 on page 590).

13.9 Remarks

Exercise 13.7 is based on a result appearing in Benoit and Krishna [1985]. Exercise 13.34 is based on Rubinstein [1982]. Exercise 13.38 is based on Neyman [1985]. Exercise 13.50 is based on Liggett and Lippman [1969]. Exercise 13.51 is based on Example H.1 in Filar and Vrieze [1997]. The game appearing in Exercise 13.53 is known as the Big Match. It was first described in Gillette [1957], and was extensively studied in Blackwell and Ferguson [1968].

A review on repeated games with complete information can be found in Sorin [1992]. For a presentation of repeated games with incomplete information see Aumann and Maschler [1995] and Sorin [2002]. For a presentation of repeated games with private monitoring see Mailath and Samuelson [2006]. More information on Tauberian Theorems, of which Theorem 13.31 (page 551) is an example, can be found in Korevaar [2004].

The authors thank Abraham Neyman for clarifications provided during the composition of this chapter.

13.10 Exercises

13.1 Compute the number of pure strategies a player has in a T-stage game with n players, where the number of actions of each player i in the base game is $|S_i| = k_i$.

13.2 Artemis and Diana are avid hunters. They devote Tuesdays to their shared hobby. On Monday evening, each of them separately writes, on a slip of paper, whether or not he or she will go hunting, and whether he or she wants to be the lead hunter, or the second hunter. They then meet and each reads what the other wrote. If at least one of the two is not interested in going hunting, or both of them want the same role (lead hunter or second hunter), they do not go hunting on Tuesday. If they are both interested in a hunt, one of them wants to be the lead hunter, and the other wants to be second hunter, they do go hunting on Tuesday.

The utility of being lead hunter is 2. The utility of being second hunter is 1, and the utility of not going hunting is 0. Answer the following questions for this situation of repeated interaction:

(a) Write down the base game for this situation.
(b) Find all the equilibria of the one-stage game (the base game).
(c) Find all the equilibria of the two-stage game.

13.3 Repeat Exercise 13.2 for the following situation. Mark and Jim are neighbors, and are employed in the same place of work. They start work at the same hour every day, but their working day ends at different hours. Each has the option of going to work by train, or by bus. Every morning, each of them decides the mode of transportion by which he will get to work that day. Each of them "gains" 5 when they travel to work together, and each "gains" 0 if they travel by different modes of transportation. Taking the bus costs 1, and taking the train costs 2. Mark enjoys a 50% reduction on train tickets. The utility each of them receives is the difference between what he gains during the ride to work, and the cost of the ticket. For example, Jim's utility from taking the bus with Mark is 4.

13.4 Repeat Exercise 13.2 for the following situation. There are two pubs in a neighborhood. Three friends, Andrew, Mike, and Ron, like to cap off their working days with a beer at the pub. Each of them gains a utility of 2 when drinking with only one other friend, a utility of 1 when the three drink together, and a utility of 0 when drinking alone. Every day each of them independently decides which of the two pubs in the neighborhood he will go to, for a drink.

13.5 Prove or disprove the following claim: let τ be a strategy vector in Γ_T where, for each history h, the mixed action vector $\tau(h)$ is an equilibrium of the base game Γ. Then τ is an equilibrium of the game Γ_T. Compare this result with Theorem 13.6 (page 528), where the equilibrium of the base game that is played at any stage is independent of the history.

13.6 Prove that at every equilibrium of the T-stage Prisoner's Dilemma, both players play D in every stage.

13.7 Let $\Gamma = (N, (S_i)_{i \in N}, (u_i)_{i \in N})$ be a game in strategic form that has a unique equilibrium, and let Γ_T be the T-stage repeated game corresponding to Γ. Prove that Γ_T has a unique subgame perfect equilibrium. Is it possible for Γ_T to have an additional Nash equilibrium? Justify your answer.

13.8 Prove that the minmax value of each player i in the T-stage repeated game is equal to his minmax value $\overline{v}_i$ in the base game.

13.9 In the following two-player zero-sum game (see Figure 13.15), find the value of the T-stage repeated game, and the optimal strategies of the two players for every $T \in \mathbb{N}$. What is the limit of the values of the T-stage games, as T goes to infinity? Player I's set of actions is $A_{\mathrm{I}} = \{T, B\}$, and Player II's set of actions is $A_{\mathrm{II}} = \{L, R\}$.

- If the players choose the pair of actions (T, L), Player II pays Player I the sum of \$1, and the players play the repeated game in Figure 13.15(A).
- If the players choose the pair of actions (T, R), Player II pays Player I the sum of \$4, and the players play the repeated game in Figure 13.15(B).
- If the players choose the pair of actions (B, L), Player II pays Player I the sum of \$2, and the players play the repeated game in Figure 13.15(B).
- If the players choose the pair of actions (B, R), Player II pays Player I the sum of \$0, and the players play the repeated game in Figure 13.15(A).

		Player II	
		L	*R*
Player I	*T*	4	0
	B	1	5

A

		Player II	
		L	*R*
Player I	*T*	1	0
	B	0	1

B

Figure 13.15 The payoff matrix of the game in Exercise 13.9

13.10 In this exercise, we will prove that the payoff received by a player in a T-stage repeated game is a linear function of the probabilities under which he chooses his pure strategies.

(a) Let $\tau_i^1, \ldots, \tau_i^L$ be all the pure strategies of player i in the T-stage repeated game. Prove that for every behavior strategy τ_i of player i (see Definition 13.3 on page 525) in the repeated game, there exist nonnegative numbers $\alpha_1, \ldots, \alpha_L$, whose sum is 1, such that for each strategy vector τ_{-i} of the other players,

$$\gamma_i^T(\tau_i, \tau_{-i}) = \sum_{l=1}^{L} \alpha_l \gamma_i^T(\tau_i^l, \tau_{-i}). \tag{13.96}$$

(b) What are the coefficients $(\alpha_l)_{l=1}^L$?

13.11 Consider the following base game.

Player II

Player I

	L	*R*
T	− 1, 3	4, 0
B	1, −1	0, 2

What is the limit of the average payoffs in the infinitely repeated game corresponding to this game, when the players implement the following strategies?

(a) In even stages, Player I plays T, and in odd stages he plays B. In stages divisible by 3 Player II plays L, and in all other stages he plays R.

(b) In even stages, Player I plays T, and in odd stages he plays B. Player II plays as follows. In the first stage he plays L. At any other stage he plays R if Player I played T in the previous stage; otherwise he plays the mixed action $[\frac{1}{4}(L), \frac{3}{4}(R)]$ in the current stage.

(c) Player I plays $[\frac{2}{3}(T), \frac{1}{3}(B)]$ in every stage. Player II plays as follows. In the first stage he plays L. At any other stage he plays R if Player I played T in the previous stage; otherwise he plays the mixed action $[\frac{1}{4}(L), \frac{3}{4}(R)]$ in the current stage.

13.12 For each of the infinitely repeated games corresponding to the following base games, plot on the same graph in $\mathbb{R}^2$ the sets F and $F \cap V$, where the x-axis represents Player I's payoff, and the y-axis represents Player II's payoff.

Player II

Player I

	L	*R*
T	−1, 1	1, −1
B	1, −1	−1, 1

Game A

Player II

Player I

	L	*R*
T	6, 6	2, 7
B	7, 2	0, 0

Game B

Player II

Player I

	L	*C*	*R*
T	0, 0	2, 4	4, 2
M	4, 2	0, 0	2, 4
B	2, 4	4, 2	0, 0

Game C

Player II

Player I

	L	*R*
T	4, 2	2, 3
B	1, 0	0, 1

Game D

13.13 Prove Theorem 13.11 on page 534: for every $K \in \mathbb{N}$ and every vector $x \in F$ there are nonnegative integers $(k_s)_{s \in S}$ summing to K satisfying

$$\left\| \sum_{s \in S} \frac{k_s}{K} u(s) - x \right\|_\infty \leq \frac{M \times |S|}{K}. \tag{13.97}$$

13.14 Suppose that each player i in a two-player zero-sum base game Γ has a unique optimal mixed strategy x_i. Prove that in the T-stage repeated game Γ_T, at each equilibrium in behavior strategies, at each stage each player implements the mixed strategy x_i.

13.15 In the 1,000,000-stage repeated game of the following base game, describe an equilibrium whose payoff is within 0.01 of (5, 6), and an equilibrium whose payoff is within 0.01 of (4, 3).

Player II

		L	R
Player I	T	6, 6	2, 7
	B	7, 2	0, 0

13.16 Consider the infinitely repeated game of the following base game.

Player II

		L	R
Player I	T	4, 6	2, 8
	B	7, 3	1, 0

Suppose that, in this game, Player I implements the following strategy σ_{I}. In the first stage, he plays the mixed action $[\frac{2}{3}(T), \frac{1}{3}(B)]$. In every stage $t > 1$, he plays a mixed action that is determined by the action that Player II played in the previous stage: if in stage t Player II played L, then in stage $t+1$, Player I plays the mixed action $[\frac{1}{2}(T), \frac{1}{2}(B)]$, while if in stage t Player II played R, then in stage $t+1$ Player I plays the mixed action $[\frac{3}{4}(T), \frac{1}{4}(B)]$.

Player II is considering which of the following four strategies to implement: (a) play L in every stage, (b) play R in every stage, (c) play L in odd stages, and R in even stages, (d) play R in odd stages, and L in even stages.

What is the limit of the average payoffs of each of the players when Player I implements strategy σ_{I} and Player II implements each of the above four strategies?

13.17 For the base game in Exercise 13.15 describe an equilibrium of the infinitely repeated game that yields a payoff of $(4\frac{1}{3}, 2\frac{1}{3})$.

13.18 For each of the following base games determine whether or not (2, 1) is an equilibrium payoff of the corresponding infinitely repeated game. If it is an equilibrium payoff, describe an equilibrium leading to that payoff. If not, justify your answer. In these games, Player I is the row player and Player II is the column player.

	L	R
T	0, 0	2, 2
B	1, 1	0, 0

Game A

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

Game B

	L	R
T	1, 3	4, 0
B	3, 0	1, 1

Game C

	L	M	R
T	0, 1	1, 0	3, 1
B	3, 1	0, 2	0, 3

Game D

13.19 In the infinitely repeated game of the following base game, describe an equilibrium leading to the payoff $(2, 3\frac{2}{3})$.

		Player II	
		L	R
Player I	T	1, 4	2, 5
	B	0, 1	3, 2

13.20 In the following three-player base game Γ Player I chooses the row (T or B), Player II chooses the column (L or R), and Player III chooses the matrix (W or E). Describe an equilibrium in the infinitely repeated game, based on Γ, for which the resulting payoff is $(2, 1, 2\frac{1}{2})$.

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

W

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

E

13.21 One of the payoffs in the following base game is a parameter labeled x. For every $x \in [0, 1]$ find the set of equilibrium payoffs in the infinitely repeated game based on this game.

	L	R
T	1, 1	0, 2
B	0, 1	$x, \frac{3}{2}$

13.22 Find an example of an infinitely repeated game, and a strategy vector τ in this game satisfying (a) τ is an equilibrium for every finite game Γ_T with a corresponding payoff of $\gamma^T(\tau)$ and (b) the limit $\lim_{T\to\infty} \gamma^T(\tau)$ does not exist.

13.23 Prove the Folk Theorem for infinitely repeated games (Theorem 13.17 on page 539).
Guidance: For each $K \in \mathbb{N}$, approximate x by a weighted average of vectors in the payoff matrix, with weights that are nonnegative and rational, with denominator

K. Construct a strategy vector in which the players play in blocks, such that the length of the K-th block is K stages, and in the K-th block the players play in such a way that the average of the payoffs is approximately x. If at a certain stage, a player deviates from the action he is supposed to play at that stage, he is punished from the next stage onwards by a punishment strategy.

13.24 Prove directly that the statement of Theorem 13.19 (page 541) holds with respect to the strategy vector τ^* defined in Example 13.1 (page 540): τ^* is an ε-equilibrium of the T-stage game, for every T sufficiently large.

13.25 Prove the strong formulation of the Folk Theorem for infinitely repeated games (Theorem 13.19 on page 541).

13.26 In the game in Example 13.20, prove that for every $T \in \mathbb{N}$, Player I has a strategy in Γ_T yielding the payoff $\frac{T-\lceil\sqrt{T}\rceil}{T}$ when Player II uses the strategy τ^*_{II} defined in the example.

13.27 Let N be a set of players, and let $(S_i)_{i\in N}$ be finite sets of actions of the players. Let $u : S \to \mathbb{R}^N$ and $u' : S \to \mathbb{R}^N$ be two payoff functions. Consider a variation of the repeated game, in which in odd stages the payoff function is u, and in even stages the payoff function is u'.

(a) Write the analogous theorem to Theorem 13.9 in this model.
(b) Write the analogous theorem to Theorem 13.17 in this model.

13.28 Repeat Exercise 13.27, under the following variation of the game: in each stage, one of the payoff functions is chosen randomly (each payoff function is chosen with probability $\frac{1}{2}$, independently of the payoff functions and the actions of the players in previous stages), and the players are informed of the chosen payoff functions before they choose their actions in each stage.

13.29 Repeat Exercise 13.28 for the case where the players are not informed of the payoff function chosen in each stage.

13.30 Repeat Exercise 13.28 for the case where only Player 1 is informed of the payoff function chosen in each stage (with the other players not informed of the chosen payoff function).

13.31 Prove that for every discount factor $\lambda \in [0, 1)$, the minmax value of each player i in the λ-discounted game Γ_λ is equal to his minmax value in the base game Γ.

13.32 Compute the λ-discounted payoff in each of the three cases (a), (b), (c) of Exercise 13.11.

13.33 **Cartel game** A cartel is an association of players who coordinate their actions in order to attain better results than the players could attain if they acted individually. In this exercise we will show that players can indeed profit by forming a cartel, and check whether a cartel can be stable.

Consider the following Cournot competition (see Example 4.23 on page 99): there are n luxury car manufacturers. The manufacturing cost of each car is

\$100,000 (for each manufacturer) and the consumer price of each such car is \$200,000 $- \sum_{i=1}^{n} x_i$, where x_i is the number of cars manufactured annually by manufacturer i. For computational ease, we assume below that x_i can be any nonnegative real number (not necessarily an integer). Answer the following questions:

(a) Describe the situation as a strategic-form game, where a pure strategy of each manufacturer is the number of cars he manufactures annually.
(b) Prove that this game has a unique symmetric equilibrium (that is, an equilibrium x in which $x_i = x_j$ for all i and j), at which $x_i = \frac{100{,}000}{n+1}$.
(c) Suppose that the manufacturers decide to form a cartel, and to determine jointly the number of cars that each of them will manufacture, in order to maximize the profit of each of them. Prove that to maximize this profit, the manufacturers need to manufacture collectively 50,000 cars; hence if they divide this number equally between them they will each manufacture $\frac{50{,}000}{n}$ cars. In other words, the cartel limits the number of cars manufactured by each member to a number that is lower than the number of cars manufactured at equilibrium (assuming that $n > 1$). Show that despite the lower manufacturing numbers, the profit of each manufacturer under the cartel's quotas is higher than his profit at the equilibrium strategy.
(d) Consider next the discounted repeated game of the above-described base game. Is the strategy vector at which each manufacturer manufactures $\frac{50{,}000}{n}$ cars in each stage an equilibrium of Γ_λ, for every $\lambda \in [0, 1)$? Justify your answer.
(e) For each manufacturer i define a strategy τ_i as follows:
  - In the first stage, manufacture $\frac{50{,}000}{n}$ cars.
  - For each $t > 1$, the number of cars to manufacture in stage t is determined as follows:
    - if in each of the previous stages every manufacturer manufactured $\frac{50{,}000}{n}$ cars, manufacture $\frac{50{,}000}{n}$ cars in stage t;
    - otherwise, manufacture $\frac{100{,}000}{n+1}$ cars in stage t.

  For which value of n, and which discount factor λ, is the strategy vector $\tau = (\tau_i)_{i=1}^{n}$ an equilibrium of the game Γ_λ? What can we conclude regarding the stability of cartels, given these results?
(f) Are there similarities between the repeated Prisoner's Dilemma (see Example 13.1 on page 521) and the cartel game of this exercise? If so, what are they? Which equilibria of the repeated Prisoner's Dilemma correspond to the equilibria described in items (b) and (e) of this exercise?

13.34 Alternating offers game Barack and Joe can together implement a project that will jointly yield them a profit \$100. How should they divide this sum of money between them? They decide to implement the following mechanism: Barack will offer Joe a split of $(x, 100 - x)$, where x is a number in the interval $[0, 100]$, signifying the amount of money that Barack will receive under this offer. Joe may accept or reject this offer. If he accepts, this will be the final split. If he rejects the offer, the next day he proposes a counteroffer $(y, 100 - y)$, where y is a number in the interval [0, 100], signifying the amount of money that Barack will receive, under this offer. Barack may accept or reject this offer. If he accepts, this will be the final split. If he rejects the offer, the next day he proposes a counteroffer, and so on. Every delay in implementing the project reduces the profit they will receive: if the two of them agree on a division of the money $(x, 100 - x)$ on day n, Barack's payoff is $\beta^{n-1} \times x$, and Joe's payoff is $\beta^{n-1} \times (100 - x)$, where $\beta \in (0, 1)$ is the discount factor in the game (in other words, $100(\frac{1}{\beta} - 1)$ is the daily interest rate in the game).

Depict this situation as an extensive-form game, and find all the subgame perfect equilibria of the game.

13.35 In the two-player zero-sum game in Exercise 13.9, find the value of the discounted game, and the optimal strategy of both players for any discount factor $\lambda \in [0, 1)$. What is the limit of the discounted values, as the discount factor converges to 1? Is the limit equal to the limit you computed in Exercise 13.9 for the values of the T-stage game?

13.36 Find an example of a repeated game, and a strategy vector τ, such that (a) τ is an equilibrium of the discounted game for every $\lambda \in [0, 1)$, and (b) the limit $\lim_{\lambda \to 1} \gamma^{\lambda}(\tau)$ does not exist.

13.37 Suppose two players are playing the repeated Prisoner's Dilemma. Prove that if the discount factor λ is sufficiently close to 1, the strategy vector at which the players implement the grim-trigger strategy, i.e., every player plays C as long as the other player plays C, and otherwise plays D, is a λ-discounted equilibrium.

13.38 A strategy in an infinitely repeated game has *recall k* if the action a player chooses in stage t depends only on the actions that were played in stages $t - 1, t - 2, \ldots, t - k$ (and is independent of the actions played in earlier stages, and of the number of the stages t). Formally, a strategy τ_i has recall k if for every $t, \widehat{t} \geq k$ we have $\tau_i(a^1, a^2, \ldots, a^{t-1}) = \tau_i(\widehat{a}^1, \widehat{a}^2, \ldots, \widehat{a}^{\widehat{t}-1})$ whenever $(a^{t-k}, a^{t-k+1}, \cdots, a^{t-1}) = (\widehat{a}^{\widehat{t}-k}, \widehat{a}^{\widehat{t}-k+1}, \cdots, \widehat{a}^{\widehat{t}-1})$.

(a) How many pure strategies of recall k has each player got?
(b) Can the grim-trigger strategy be implemented by a pure strategy with recall k? Justify your answer.
(c) Prove that in the T-stage repeated Prisoner's Dilemma, when the players are limited to playing only strategies with recall k (where $k + 1 < T$), (3, 3) is an equilibrium payoff.
(d) For which triples k, l, and T is (3, 3) an equilibrium payoff in the T-stage repeated Prisoner's Dilemma, where Player I is limited to strategies with recall k, and Player II is limited to strategies with recall l?

13.39 Suppose two players are playing the repeated Prisoner's Dilemma with an unknown number of stages; after each stage, a lottery is conducted, such that with probability $1 - \beta$ the game ends with no further stages conducted, and with probability β the game continues to another stage, where $\beta \in [0, 1)$ is a given real number. Each player's goal is to maximize the sum total of payoffs received over all the stages of the game.

Prove that if β is sufficiently close to 1, the strategy vector in which at the first stage every player plays C, and in each subsequent stage each player plays C if the other player played C in the previous stage, and he plays D otherwise, is an equilibrium. This strategy is called the Tit-for-Tat strategy.

13.40 In this exercise, we will show that in a discounted two-player zero-sum game in which the discount factors of the two players are different from each other, the payoff to each player at every equilibrium is the value of the base game. Consider the two-player zero-sum repeated game based on the following base game.

	L	R
T	$-1, 1$	$1, -1$
B	$1, -1$	$-1, 1$

Assume that the discount factor of Player I is λ throughout this exercise (except in section (j)), and that the discount factor of Player II is λ^2, where $\lambda \in [0, 1)$. Answer the following questions:

(a) What is the value in mixed strategies v of the base game?

(b) What is the discounted payoff to each player under the following pair of strategies, as a function of the parameter $t_0 \in \mathbb{N}$:
  - Player I plays T in each stage of the game.
  - Player II plays L in the first t_0 stages, and always R afterwards.

(c) Find t_0 such that the sum of the payoffs of the two players is maximized. What is the sum of the payoffs for this t_0? In the solution here, assume that t_0 may be any nonnegative real number.

(d) Prove that the pair of strategies in which Player I plays the mixed action $[\frac{1}{2}(T), \frac{1}{2}(B)]$ at each stage, and Player II plays the mixed action $[\frac{1}{2}(L), \frac{1}{2}(R)]$ at each stage, is an equilibrium in this discounted game.

Let $\tau^* = (\tau^*_{\mathrm{I}}, \tau^*_{\mathrm{II}})$ be any equilibrium of this discounted game. For $t_0 \in \mathbb{N}$, denote by $A(\lambda, t_0)$ the λ-discounted payoff under strategy vector τ^* starting from stage t_0:

$$A(\lambda, t_0) = (1 - \lambda) \sum_{t=t_0}^{\infty} \mathbf{E}_{\tau^*}[u^t] \lambda^{t-t_0-1}. \tag{13.98}$$

(e) Prove that for every $t_0 \in \mathbb{N}$, the following holds: $A(\lambda, t_0) \geq v$ and $A(\lambda^2, t_0) \leq v$.

(f) Prove that for every $t_0 \in \mathbb{N}$ and every $\lambda \in [0, 1)$, the following holds:

$$A(\lambda, t_0) = A(\lambda^2, t_0) + \sum_{k=1}^{\infty} \lambda^k (1 - \lambda) A(\lambda^2, t_0 + k). \tag{13.99}$$

(g) Deduce from the last two items that $A(\lambda, t_0) = v$ for every $t_0 \in \mathbb{N}$, and from this further deduce that $\mathbf{E}_{\tau^*}[u^t] = 0$ for every $t \in \mathbb{N}$.

(h) Prove that at each equilibrium of this discounted game, the discounted payoff of each player is v.

(i) Does the result of item (c) contradict the result of item (h)? Explain.

(j) Generalize the result of item (h) to any discounted game and any pair of discount factors: if $\tau^* = (\tau_{\mathrm{I}}^*, \tau_{\mathrm{II}}^*)$ is an equilibrium of a two-player zero-sum game in which the discount factor of Player I is λ_{I} and the discount factor of Player II is λ_{II}, then $A(\lambda_i, t_0) = v$ for $i \in \{\mathrm{I}, \mathrm{II}\}$, for every $t_0 \in \mathbb{N}$, where v is the value in mixed strategies of the base game. In particular, at any equilibrium, the discounted payoff of each player (at his discount factor) is the value of the base game.

13.41 Prove that the condition in Theorem 13.9 (page 531) implies the condition in Theorem 13.24 (page 545): if for every player i there exists an equilibrium $\beta(i)$ in the base game for which $\beta_i(i) > \overline{v}_i$, then there exists a vector $\widehat{x} \in F \cap V$ satisfying $x_i > \overline{v}_i$ for every $i \in N$.

13.42 Prove the Folk Theorem for discounted games (Theorem 13.24 on page 545).

13.43 Show (by finding appropriate strategy vectors) that the payoff vectors mentioned in Exercises 13.19 and 13.20 are payoffs of uniform equilibria for discounted games.

13.44 Prove the Folk Theorem for uniform equilibrium in discounted games (Theorem 13.27 on page 549).

13.45 In this exercise, we define the uniform value of two-player zero-sum games. Let Γ be a two-player zero-sum base game. The real number v is called the *uniform value* (for the finite games $(\Gamma_T)_{T \in \mathbb{N}}$) if for each $\varepsilon > 0$ there exist strategies τ_{I}^* of Player I and τ_{II}^* of Player II in Γ_∞, and an integer T_0, such that the following condition is satisfied: for each $T \geq T_0$, and each pair of strategies $(\tau_{\mathrm{I}}, \tau_{\mathrm{II}})$ in Γ_T,

$$\gamma^T(\tau_{\mathrm{I}}, \tau_{\mathrm{II}}^*) \leq v + \varepsilon, \text{ and } \gamma^T(\tau_{\mathrm{I}}^*, \tau_{\mathrm{II}}) \geq v - \varepsilon. \tag{13.100}$$

Prove that the uniform value for finite games equals the value of the base game.

13.46 Repeat Exercise 13.43 for uniform ε-equilibria for finite games, for every $\varepsilon > 0$.

13.47 Let E_T be the set of equilibrium payoffs of a T-stage repeated game Γ_T.

(a) Prove that $E_T \subseteq E_{kT}$ for every $T \in \mathbb{N}$ and for every $k \in \mathbb{N}$.

(b) Prove[8] that $\frac{T}{T+1} E_T + \frac{1}{T+1} E_1 \subseteq E_{T+1}$ for every $T \in \mathbb{N}$.

---

[8] For every pair of sets S_1 and S_2 in $\mathbb{R}^k$, and every real number α, the sets αS_1 and $S_1 + S_2$ are defined by $\alpha S_1 := \{\alpha x : x \in S_1\}$ and $S_1 + S_2 := \{x + y : x \in S_1, \quad y \in S_2\}$.

(c) For every set $S \subseteq \mathbb{R}^k$, the set $\overline{S}$ denotes the closure of S: the smallest closed set containing S. Let

$$E_\infty := \limsup_{T \to \infty} E_T = \bigcup_{T \in \mathbb{N}} \bigcap_{k \geq T} \overline{E}_k. \tag{13.101}$$

The set E_∞ is the upper limit of the sets $(E_T)_{T \in \mathbb{N}}$, and it includes all the partial limits of the sequences $(x_t)_{t \in \mathbb{N}}$, where $x_t \in E_t$ for each $t \in \mathbb{N}$.

Prove, using items (a) and (b), that $E_T \subseteq E_\infty$ for every $T \in \mathbb{N}$, and in particular that E_∞ is not empty. Furthermore, prove that for every $x \in E_\infty$ and every $\varepsilon > 0$, there exists $T_0 \in \mathbb{N}$ such that for every $T \geq T_0$ there exists $y \in E_T$ satisfying $\|x - y\|_\infty \leq \varepsilon$. In other words, the sets $(E_T)_{T \in \mathbb{N}}$ "approach" E_∞ as T goes to infinity.

13.48 Prove that in every uniform 0-equilibrium for finite games, from some stage onwards the players play an equilibrium of the base game at each stage.

13.49 Prove the Folk Theorem for uniform equilibrium in finite games (Theorem 13.29 on page 549).

13.50 In this exercise we prove the existence of a sequence $(x_t)_{t=1}^\infty$ of zeros and ones satisfying

$$\liminf_{T \to \infty} \frac{\sum_{k=1}^T x_k}{T} < \liminf_{\lambda \to 1} (1 - \lambda) \sum_{t=1}^\infty \lambda^{t-1} x_t. \tag{13.102}$$

Let $(q_t)_{t \in \mathbb{N}}$ be a sequence of natural numbers. Define a sequence $(p_t)_{t \in \mathbb{N}}$ as follows:

$$p_1 := 0, \tag{13.103}$$

$$p_t := q_1 + q_2 + \cdots + q_{t-1}. \tag{13.104}$$

Define a sequence $(x_t)_{t \in \mathbb{N}}$ as follows:

$$x_t = \begin{cases} 1 & \text{when there exists } k \text{ such that } 2p_k < t \leq 2p_k + q_k, \\ 0 & \text{otherwise.} \end{cases} \tag{13.105}$$

In words, the first q_1 elements of the sequence $(x_t)_{t \in \mathbb{N}}$ equal 1, the next q_1 elements of the sequence equal 0, the next q_2 elements of the sequence equal 1, the next q_2 elements of the sequence equal 0, and so on.

(a) Prove that $\liminf_{T \to \infty} \frac{\sum_{k=1}^T x_k}{T} = \frac{1}{2}$.
(b) Denote $A(\lambda) = (1 - \lambda) \sum_{t=1}^\infty \lambda^{t-1} x_t$. Prove that $A(\lambda) = \sum_{k=1}^\infty \lambda^{2p_k}(1 - \lambda^{q_k})$.
(c) Denote $\alpha_k = \lambda^{p_k} - \lambda^{p_{k+1}}$ for every $k \in \mathbb{N}$. Using item (b) above, prove that

$$A(\lambda) = \frac{1}{2} \left(\sum_{k=1}^\infty (\alpha_k)^2 + 1 \right). \tag{13.106}$$

(d) Let $\varepsilon \in (0, \frac{1}{4})$, and define $c := \frac{\ln(\varepsilon)}{\ln(1 - \sqrt{\varepsilon})}$. Prove that $c > 2$.

(e) Suppose that the sequence $(q_t)_{t\in\mathbb{N}}$ satisfies $q_k > \frac{2p_k}{c-2}$ for every $k \in \mathbb{N}$. Define

$$a_k := \frac{|\ln(1-\sqrt{\varepsilon})|}{q_k}, \qquad b_k := \frac{|\ln(\varepsilon)|}{2p_k}. \tag{13.107}$$

Prove that $\lim_{k\to\infty} b_k = 0$, and that for every $k \in \mathbb{N}$, (a) $b_{k+1} < b_k$ for every $k \in \mathbb{N}$, (b) $cq_k > 2p_k + 2q_k$, and (c) $a_k < b_{k+1}$.

(f) Prove, with the aid of item (e) above, that $\bigcup_{k\in\mathbb{N}}(a_k, b_k) = (0,\infty)$. Deduce that for every $\lambda \in [0,1)$ there exists $k(\lambda) \in \mathbb{N}$ satisfying $a_{k(\lambda)} \le |\ln(\lambda)| < b_{k(\lambda)}$.

(g) Using Equation (13.107), prove that $\varepsilon < \lambda^{2p_{k(\lambda)}}$, and $1-\sqrt{\varepsilon} \ge \lambda^{q_{k(\lambda)}}$. Deduce that

$$(\alpha_k)^2 = \lambda^{2p_{k(\lambda)}}(1-\lambda^{q_{k(\lambda)}})^2 > \varepsilon^2. \tag{13.108}$$

(h) Deduce, with the aid of item (c) above, that $\liminf_{\lambda\to 1} A(\lambda) \ge \frac{\varepsilon^2+1}{2}$.

(i) Deduce that Equation (13.102) holds for the sequence $(x_t)_{t\in\mathbb{N}}$ defined in item (d) above.

(j) Construct a sequence $(y_t)_{t=1}^{\infty}$ of zeros and ones satisfying

$$\limsup_{T\to\infty} \frac{\sum_{t=1}^{T} y_t}{T} > \limsup_{\lambda\to 1}(1-\lambda)\sum_{t=1}^{\infty}\lambda^{t-1}y_t. \tag{13.109}$$

Such a sequence was used in Example 13.33.

13.51 Consider the following sequence $(x_t)_{t\in\mathbb{N}}$: $1, -1, 2, -2, 3, -3, \ldots$, i.e., $x_{2t} = -t$, and $x_{2t-1} = t$ for every $t \in \mathbb{N}$. Compute $\limsup_{T\to\infty} \frac{\sum_{k=1}^{T} x_k}{T}$, $\liminf_{T\to\infty} \frac{\sum_{k=1}^{T} x_k}{T}$ and $\lim_{\lambda\to 1}(1-\lambda)\sum_{t=1}^{\infty}\lambda^{t-1}x_t$.

13.52 Prove that the strategy vector $(\sigma_{\mathrm{I}}^{A}, \sigma_{\mathrm{II}})$ defined in Example 13.33 (page 553) is not an equilibrium of the infinitely repeated game.

13.53 David and Tom play the following game, over T stages. In each stage David chooses a color, either red or yellow, and Tom guesses which color David chose. If Tom guesses "red," he pays David one dollar if he guessed incorrectly, and receives one dollar from David if he guessed correctly. If, however, Tom guesses "yellow," he pays David a dollar in that stage and in every subsequent stage of the game if he guessed incorrectly, and he receives a dollar from David in that stage and in every subsequent stage of the game if he guessed correctly.

Note that this is not a repeated game, because if the first time that Tom guesses "yellow" is in stage t, the payoffs in all the stages after t depend on Tom's choice in stage t.

(a) Prove that the only equilibrium payoff when $T = 1$ is $(0, 0)$.

(b) Prove that the only equilibrium payoff when $T = 2$ is $(0, 0)$.

(c) Prove that the only equilibrium payoff for every T is $(0, 0)$.

13.54 Consider the game in Exercise 13.53 with $T = \infty$. Let x and y be two numbers in the interval $[0, 1]$. Suppose that in each stage David chooses "red" with probability

x and “yellow” with probability $1 - x$, and in each stage Tom guesses “red” with probability y and “yellow” with probability $1 - y$.

(a) Compute the expected λ-discounted payoff in this infinite game as a function of x and y, for each $\lambda \in [0, 1)$.

(b) Conclude that, if the players are restricted to these i.i.d. strategies, $(0, 0)$ is a λ-discounted equilibrium payoff, for each $\lambda \in [0, 1)$. What are the corresponding equilibrium strategies?