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 $N = \{1, 2, \ldots, n\}$ 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

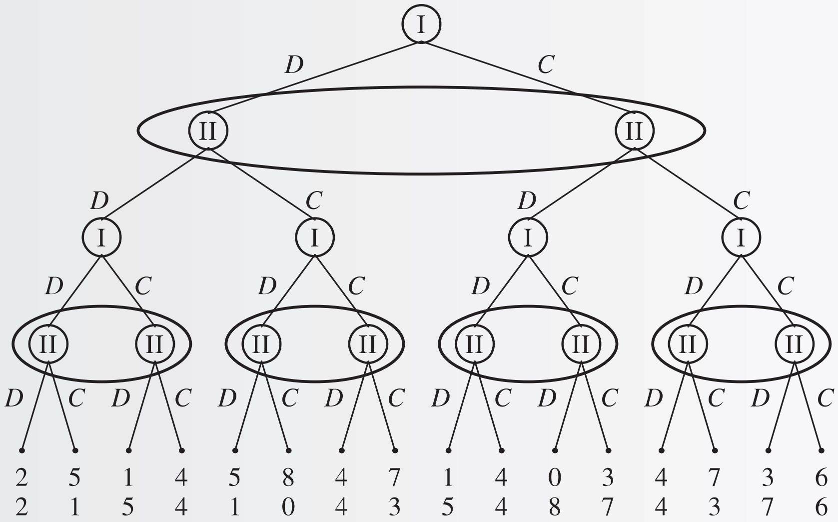

Figure 13.2 The two-stage Prisoner's Dilemma, represented as an extensive-form game

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 $(s^1, s^2, \ldots, s^t)$, 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 $t \in \{0, 1, \ldots, T-1\}$.

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.

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

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 $\tau = (\tau_1, \tau_2, \dots, \tau_n)$ 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 $\tau_i(s_i \mid s^1, \dots, s^{t-1})$ the conditional probability that player i plays action s_i in stage t, given that the players have played $(s^1, \dots, s^{t-1})$ 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}.

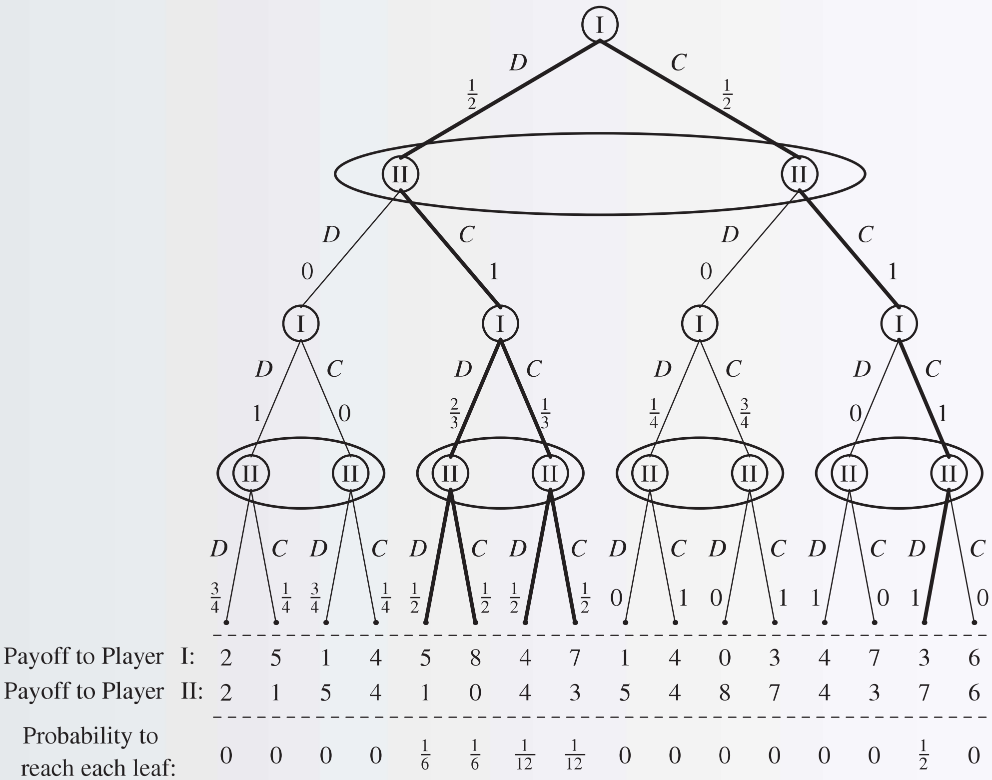

Figure 13.5 The probabilities attached to each play of the game, under the strategies $(\tau_{\mathrm{I}}, \tau_{\mathrm{II}})$ ◀

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 $(s^1, s^2, \dots, s^T)$ 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 $s^1 = (s^1_1, \dots, s^1_n) \in S$, define

$$\mathbf{P}_\tau(s^1) = \tau_1\big(s^1_1\big) \times \tau_2\big(s^1_2\big) \times \cdots \times \tau_n\big(s^1_n\big). \tag{13.4}$$

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$(s^1, s^2, \ldots, s^t) \in S^t$, define by induction

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

This means that the probability that under τ the players play the action vector $s^1, s^2, \ldots, s^t$ in the first t stages is the probability that the players play $s^1, s^2, \ldots, s^{t-1}$ 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 $s^1, s^2, \ldots, s^{t-1}$ 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 $u^t = (u_1^t, \ldots, u_n^t)$. Then, during the course of a play of the game, player i receives the sequence of payoffs $(u_i^1, u_i^2, \ldots, u_i^T)$. 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 $\mathbf{E}_\tau[f]$:

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

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

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

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

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

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

The last equality follows from the fact that the sum total of the probabilities of all $(t-1)$-stage histories is 1, and therefore $\mathbf{E}_{\tau_i,\tau^*_{-i}}[u_i^t] \leq u_i(\sigma^t)$. Averaging over the T stages of the game shows that $\gamma_i^T(\tau_i,\tau^*_{-i}) \leq \gamma_i^T(\tau^*)$, which is what we wanted to show. Since $\gamma_i^T(\tau_i,\tau^*_{-i}) \leq \gamma_i^T(\tau^*)$ 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 $\gamma_i^T(\tau^*)\geq\overline{v}_i$ 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 $\gamma_i^T(\tau_i^*,\tau_{-i})\geq\overline{v}_i$.

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 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