

# 15

## Repeated Games and the Folk Theorem

In the previous chapter we saw what can happen when players play different games over time, and more importantly, how they can use conditional strategies to support behavior that would not be optimal in a static setting. Namely, when there is a future, there may be the possibility of using reward-and-punishment strategies to sustain behavior that is better than the short-run self-interested behavior.

A special case of multistage games has received a lot of attention over the years, and that is the case of *repeated games*. A *repeated game* is a special case of a multistage game in which the *same stage game* is being played at every stage. These games have been studied for two primary reasons. First, it is natural to view the repeated game as a depiction of many realistic settings such as firms competing in the same market day after day, politicians engaging in pork-barrelling session after session, and workers in a team production line who perform some joint task day after day.

Second, it turns out that having a game that is repeated time and time again will often result in a very appealing mathematical structure that makes the analysis somewhat simple and elegant. As rigorous social scientists, elegance is an extra bonus when it appears. In this section we will get a glimpse of the analysis of these

games, and see the extreme limits of what we can do with reward-and-punishment strategies.

## 15.1 Finitely Repeated Games

A finitely repeated game is, as its name suggests, a stage game that is repeated a finite number of times. Given that we have defined a multistage game in the previous section, we can easily define a finitely repeated game as follows:

**Definition 28** *Given a stage-game  $G$ ,  $G(T, \delta)$  denotes the **finitely repeated game** in which the stage game  $G$  is played  $T$  consecutive times, and  $\delta$  is the **common discount factor**.*

As an example, consider the 2-stage repeated game in which the following game is repeated twice with a discount factor of  $\delta = 1$ :

	$M$	$F$	$R$
$M$	4, 4	$\overline{-1, 5}$	0, 0
$F$	$\overline{5, -1}$	$\overline{1, 1}$	0, 0
$R$	0, 0	0, 0	$\overline{3, 3}$

As the matrix shows, there are two pure strategy Nash equilibria, and they are Pareto ranked, that is,  $(R, R)$  is better than  $(F, F)$  for both players. Thus, we have a “carrot” and a “stick” to use in an attempt to discipline first period behavior as was shown in the analysis of multistage games. This implies that for a high enough discount factors we may be able to support SPE behavior in the first stage that is not a Nash equilibrium of the first period stage game.

Indeed, you should be able to convince yourself that for a discount factor  $\delta \geq \frac{1}{2}$  the following strategy played by each player will be a SPE:

- Play  $M$  in the first stage. If  $(M, M)$  was played then play  $R$  in the second stage, and if  $(M, M)$  was not played then play  $F$  in the second stage.<sup>1</sup>

---

<sup>1</sup>In case you had trouble convincing yourself, here is the argument: In the second stage the players are clearly playing a Nash equilibrium. To check that this is a SPE we need to check that players would not want to deviate from  $M$  in the first stage of the game. Consider player 1 and observe that,  $u_1(M, s_2) = 4 + 3 \cdot \delta$  and  $u_1(F, s_2) = 5 + 1 \cdot \delta$ , which implies that  $M$  is a best response if and only if  $-1 \geq -2\delta$ , or,  $\delta \geq \frac{1}{2}$ .

In fact, the players ability to provide “reward and punishment” strategies in this game are very similar to that in the Prisoner-Revenge game we analyzed in the previous section. Here, both players can lose 2 in the second period by deviating away from the proposed path of play, and the temptation to deviate from  $(M, M)$  is from the added payoff of 1 in the first period. The difference is that in this example the same game is repeated twice, and it’s multiplicity of equilibria is giving us the leverage to impose conditional strategies of the reward-and-punishment type.

The conclusion is that we need *multiple continuation Nash equilibria* to be able to support behavior that is not a Nash equilibrium of the stage game in earlier stages. This immediately implies the following result (stated for multistage games, which immediately applies to any repeated game):

**Proposition 11** *If a finite multistage game consists of stage games that each have a unique Nash equilibrium, then the multistage game has a unique Subgame perfect equilibrium.*

The proof of this simple observation follows immediately from backward induction. In the last stage the players must play the unique Nash equilibrium of that game. In the stage before last they cannot condition future (last stage) behavior on current stage outcomes, so here again they must play the unique Nash equilibrium of the stage game continues by the Nash equilibrium of the last stage game, and this induction argument continues till the first stage of the game.

What are the consequences of this proposition? It immediately applies to a finitely repeated game like the Prisoner’s dilemma or like any of the simultaneous move market games (Cournot and Bertrand competition) that we analyzed. If such a game is played  $T$  times in a row as a finitely repeated game, then there will be a *unique SPE* that is the repetition of the static non-cooperative Nash equilibrium play in every stage.

This result may seem rather disturbing, but it follows from the essence of credibility and sequential rationality. If two players play the Prisoner’s dilemma for, say, 500 times in a row, then we would be tempted to think that they can use the future to discipline behavior and try to cooperate by playing  $M$ , and deviations would be punished by a move to  $F$ . However, the “unravelling” that proves Proposition X.X will apply to this finitely repeated Prisoner’s dilemma as follows: the

players must play fink in the last stage since it is the unique Nash equilibrium. In the stage before last we cannot provide reward and punishment incentives, so here again the players must play fink. This argument will continue, thus proving that we cannot support a Subgame Perfect equilibrium in which the players can cooperate in early stages of the game.

For many readers this argument is rather counter-intuitive at first, since we realize that the players have a long future ahead of them, and this should intuitively be enough to provide the players with tools, i.e., equilibrium strategies, that offer them incentives to cooperate. The twist is that the provision of such incentives rely crucially on the ability to choose reward and punishment continuation equilibrium strategies, and these strategies themselves rely on multiple equilibria in the continuation of play. As we will now see, taking the idea of a future in a somewhat different interpretation will restore our intuition, but will open the discussion in a rather surprising way.

## 15.2 Infinitely Repeated Games

The problem we identified with the finitely repeated Prisoner's dilemma was due to the *finiteness of the game*. Once a game is finite, this implies that there is a last period  $T$  from which the "Nash unravelling" of the static Nash equilibrium followed. The players must play the unique Nash equilibrium in the last stage, implying that no "reward-and-punishment" strategies are available. This in turn dictates their behavior in the stage before hand, and the game unravels accordingly.

What would happen if we eliminate this problem from its root by assuming that the game does not have a final period? That is, whatever stage the players are at, and whatever they had played, there is always a "long" future ahead of the players upon which they can condition their behavior. As we will now see, this will indeed cause the ability to support play that is not a static Nash equilibrium of the stage game. Most interestingly, we have a lot of freedom in supporting play that is not a static Nash equilibrium. First, we need to define payoffs for an infinitely repeated game.

## 15.2.1 Payoffs

Let  $G$  be the stage game, and denote by  $G(\delta)$  the infinitely repeated game with discount factor  $\delta$ . The natural extension of the present value notion from the previous section is,

**Definition 29** Given the discount factor  $\delta < 1$ , the **present value of an infinite sequence of payoffs**  $\{u_i^t\}_{t=1}^\infty$  for player  $i$  is:

$$\begin{aligned} u_i &= u_i^1 + \delta u_i^2 + \delta^2 u_i^3 + \cdots + \delta^{t-1} u_i^t + \cdots \\ &= \sum_{t=1}^{\infty} \delta^{t-1} u_i^t . \end{aligned}$$

Notice that unlike the case of a finitely repeated game (or finite multistage game) we *must have* a discount factor that is less than 1 for this expression to be well-defined. This follows because for  $\delta = 1$  this sum may not be well defined if there are infinitely many positive stage-payoffs. In such a case this sum may equal infinity. If the payoffs  $u_i^t$  are *bounded*, i.e., all stage payoffs  $u_i^t$  are less than some (potentially large) number, then with a discount factor  $\delta < 1$  this sum is well defined.<sup>2</sup>

As we argued for multistage games, this discounting is a natural economic assumption for a sequence of payoffs. The question is, however, whether it is reasonable to assume that players will engage in a infinite sequence of play. Arguably, a sensible answer is that they won't. Fortunately, we can give a reasonable motivation for this idea if instead of using the interest rate interpretation of discounting, we consider the alternative interpretation of an uncertain future.

Imagine that the players are playing this stage game today, and with some probability  $\delta < 1$  they will play the game again tomorrow, and with probability  $1 - \delta$  they end the relationship. If the relationship continues then again with probability  $\delta < 1$  they continue for another period, and so on.

Now imagine that in the event that these players would continue playing the game, the payoffs of player  $i$  are given by the sequence  $\{u_i^t\}_{t=1}^\infty$ . If there is no

---

<sup>2</sup>To see this, assume that there exists some bound  $b > 0$  such that  $u_i^t < b$  for all  $t$ . First observe that  $u_i = \sum_{t=1}^{\infty} \delta^{t-1} u_i^t < \sum_{t=1}^{\infty} \delta^{t-1} b$ . We are left to show that  $B = \sum_{t=1}^{\infty} \delta^{t-1} b$  is finite. We can write  $B(1 - \delta) = b$  because  $\lim_{t \rightarrow \infty} \delta^{t-1} b = 0$ , which in turn implies that  $B = \frac{b}{1-\delta} < \infty$ .

additional discounting,<sup>3</sup> we can think of the *expected payoff* from this sequence as,

$$\begin{aligned} Eu &= u_i^1 + \delta u_i^2 + (1 - \delta) \cdot 0 + \delta^2 u_i^3 + \delta(1 - \delta) \cdot 0 + \delta^3 u_i^4 + \delta^2(1 - \delta) \cdot 0 + \dots \\ &= \sum_{t=1}^{\infty} \delta^{t-1} u_t^i . \end{aligned}$$

That is, we can think of the present value of the payoffs as the the expected value from playing this game with this payoff sequence. There is something rather appealing about this interpretation of  $\delta < 1$  as the probability that the players do not break their ongoing relationship. Namely, what is the probability that the players will play this game infinitely many times? The answer is clearly  $\delta^\infty = 0!$  That is, the game will end in finite time with probability 1, but the *potential* future is always infinite.

To drive the point home, consider a *very high* probability of continuation, say  $\delta = 0.97$ . For this case the probability of playing for at least 100 periods is  $0.97^{100} = 0.0475$ , and for at least 500 periods is  $0.97^{500} = 2.4315 \times 10^{-7}$  (less than one in four million!). Nonetheless, this will give our players a lot of leverage as we will shortly see. Thus, we need not literally imagine that these players will play infinitely many games, but just have a chance to continue the play after every stage.

Before continuing with the analysis of infinitely repeated games, it is useful to introduce the following definition:

**Definition 30** *Given  $\delta < 1$ , the **average payoff** of an **infinite sequence of payoffs**  $\{u_i^t\}_{t=1}^{\infty}$  is:*

$$\bar{u}_i = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i^t .$$

That is, the average payoff from a sequence is a *normalization* of the net present value: we are scaling down the net present value by a factor of  $1 - \delta$ . This is

---

<sup>3</sup>If there were discounting in addition to an uncertain future, then let  $\delta < 1$  be the probability of continuation, and let  $\rho < 1$  be the discount factor. Then, the *effective discount factor* would just be  $\delta\rho$ , and the expected discounted payoff would be given by

$$\begin{aligned} Eu &= u_i^1 + \delta\rho u_i^2 + (1 - \delta) \cdot 0 + (\delta\rho)^2 u_i^3 + \delta(1 - \delta) \cdot 0 + (\delta\rho)^3 u_i^4 + \delta^2(1 - \delta) \cdot 0 + \dots \\ &= \sum_{t=1}^{\infty} (\delta\rho)^{t-1} u_t^i . \end{aligned}$$

convenient because of the following mathematical property. Consider a player that gets a finite sequence of fixed payments  $v$  in each period  $t \in \{1, 2, \dots, T\}$ . His net present value is then the sum,

$$s(v, \delta, T) = v + \delta v + \delta^2 v + \dots + \delta^{T-2} v + \delta^{T-1} v . \quad (15.1)$$

Consider this sum multiplied by  $\delta$ , that is,

$$\delta s(v, \delta, T) = \delta v + \delta^2 v + \delta^3 v + \dots + \delta^{T-1} v + \delta^T v , \quad (15.2)$$

and notice that from (15.1) and (15.2) we obtain that  $s(v, \delta) - \delta s(v, \delta) = v - \delta^T v$ , implying that

$$s(v, \delta, T) = \frac{v - \delta^T v}{1 - \delta} .$$

Now if we take the limit of this expression when  $T$  goes to infinity (the game is infinitely repeated) then we get,

$$s(v, \delta) = \lim_{T \rightarrow \infty} s(v, \delta, T) = \frac{v}{1 - \delta} .$$

How does this relate to the average payoff? If we look at the net present value of a stream of fixed payoffs  $v$ , then by definition the average payoff is,

$$\bar{v} = (1 - \delta)s(v, \delta) = v ,$$

which is very convenient. Namely, the average payoff of an infinite stream of some value  $v$  is itself  $v$ . As we will shortly see, this will help us identify what outcomes, as defined by the obtained average payoffs, can be supported as a SPE of a repeated game.

### 15.2.2 Strategies

If we think of the extensive form representation of an infinitely repeated game then we imagine a tree with a root, and it just continues from there to expand both in “length”—because of the added stages—and in “width” —because more and more information sets are created in each period. This implies that there will be an *infinite number of information sets!* Since a player’s strategy is a *complete*

*contingent plan* that tells the player what to do in each information set, it may seem that defining strategies for players in an infinitely repeated game will be quite demanding.

It turns out that there is a convenient and rather natural way to describe strategies in this setup. Notice that every information set of each player is identified by a *unique path of play* or *history*, that was played in the previous sequences. For example, if we consider the infinitely repeated Prisoner's dilemma, then in the fourth stage each player has 64 information sets ( $4^3$  since there are 4 possible outcomes in each stage). Each of these 64 information sets corresponds to a unique path of play, or history, in the first three stages. For example, the players playing  $(M, M)$  in each of the three previous stages is one such history. Any other combinations of three plays will be identified with a different and unique history.

This observation implies that there is a one-to-one relationship between information sets and *histories of play*. From now we use the word *history* to describe a particular sequence of action profiles that the players have chosen up till the stage that is under consideration (for which past play is indeed the history). To make things precise, we introduce histories, and history contingent strategies as follows:

**Definition 31** Let  $H_t$  denote the set of all possible **histories of length**  $t$ ,  $h_t \in H_t$ , and let  $H = \cup_{t=1}^{\infty} H_t$  be the set of **all possible histories** (the union over  $t$  of all the sets  $H_t$ ). A **pure strategy** for player  $i$  is a mapping  $s_i : H \rightarrow S_i$  that maps histories into actions of the stage game. (Similarly, a **behavioral strategy** of player  $i$  maps histories into random choices of actions in each stage.)

The interpretation of a strategy is basically the same as what we had for multistage games, which was clearly demonstrated by the Prisoner-Revenge game we had analyzed. Namely, if the actions in stage 2 are contingent on what happened in period 1, then this is precisely a case of history dependent strategies. Since, as we have demonstrated above, every information set is identified with a unique history, this is a complete definition of a strategy that is very intuitive to perceive.



### 15.2.3 Subgame Perfect Equilibria

Now that we have identified a concise way to write down strategies for the players in an infinitely repeated game, it is straightforward to define a SPE:

**Definition 32** *A profile of (pure) strategies  $s_i : H \rightarrow S_i$ , for all  $i \in N$ , is a **Subgame Perfect Equilibrium** if the restriction of  $(s_1(\cdot), s_2(\cdot), \dots, s_n(\cdot))$  is a Nash equilibrium in every subgame. That is, for any history of the game  $h_t$ , the continuation play dictated by  $(s_1(\cdot), s_2(\cdot), \dots, s_n(\cdot))$  is a Nash equilibrium.*

As with the case of strategies, this may at first seem like an impossible concept to implement. How could we check that a profile of strategies is a Nash equilibrium for any history, especially since each strategy is a mapping that works on any one of the infinitely many histories? As it turns out, one familiar result can be stated:

**Proposition 12** *Let  $G(\delta)$  be an infinitely repeated game, and let  $(\sigma_1^*, \sigma_2^*, \dots, \sigma_n^*)$  be a (static) Nash equilibrium strategy profile of the stage game  $G$ . Define the repeated game strategy for each player  $i$  to be the unconditional Nash strategy,  $\sigma_i^*(h) = \sigma_i^*$  for all  $h \in H$ . Then  $(s_1^*(h), \dots, s_n^*(h))$  is a Subgame Perfect equilibrium in the repeated game for any  $\delta < 1$ .*

The proof is quite straightforward, and basically mimics the ideas of the proof of proposition X.X for finite multistage games. If player  $i$  believes that his opponents' behavior is unconditional of the history then there can be no role to consider how current play affects future play. Then, if he believes that his opponents' current play coincides with their play in the static Nash equilibrium of the stage game, then by definition his best response must be to choose his part of the Nash equilibrium.

This proposition demonstrated that for the infinitely repeated Prisoner's dilemma, playing fink unconditionally in every period by each player is a SPE. Since again, the question is whether or not we can support other types of behavior as part of a SPE. In particular, can we have the players choose mum for a significant period of time in equilibrium? The answer is yes, and in many ways the intuition rests on similar arguments to the "rewards and punishments" strategies we saw in the Prisoner-Revenge game analyzed earlier.

Before discussing some general results, however, we start with an example of the infinitely repeated Prisoner's dilemma with discount factor  $\delta < 1$  and the stage

game as follows:

		Player 2	
		$M$	$F$
Player 1	$M$	$4, 4$	$-1, 5$
	$F$	$5, -1$	$1, 1$

Recall the players' *average payoffs* as defined in Definition X.X and consider the two players receiving average payoffs of  $(\bar{u}_1, \bar{u}_2) = (4, 4)$ , which would result from both playing playing  $M$  in every period. Can this be supported as a SPE? Clearly, the unconditional strategies of “play  $M$  regardless of the history” cannot be a SPE: following a deviation of player 1, player 2 will continue to play  $M$ , making the deviation profitable for player 1. This implies that if we can support  $(M, M)$  in each period then we need to “punish” deviations.

From the Prisoner-Revenge game we analyzed, it is quite clear what we need in order to support any “good” behavior: we need some “carrot” continuation equilibrium and some “stick” continuation equilibrium, and a high enough discount factor. A clear candidate for a “stick” continuation is the repeated play of the  $(F, F)$  static Nash which gives each player and average payoff of 1. What then, can be the “carrot” continuation equilibrium?

This is where things get tricky, yet in a very ingenious way. Recall that we are trying to support the play  $(M, M)$  in every period as a SPE. If we can indeed do this, it means that playing  $(M, M)$  forever *would be part of a SPE* with appropriately defined strategies. If this were possible, then continuing with  $(M, M)$  would be a natural candidate for the “carrot” continuation play. This is indeed what we will do. Assuming that  $(M, M)$  can be the “carrot”, will  $(F, F)$  be a sufficient “stick”? Consider the following strategy for each player:

- In the first stage play  $s_i^1 = M$ . For any stage  $t > 1$ , play  $s_i^t(h_{t-1}) = M$  if and only if the history is a sequence that consists only of  $(M, M)$ , that is, if and only if  $h_{t-1} = \{(M, M), (M, M), \dots, (M, M)\}$ . Otherwise, if some player ever player  $F$ , that is, if  $h_{t-1} \neq \{(M, M), (M, M), \dots, (M, M)\}$  then play  $s_i^t(h_{t-1}) = F$ .

Before checking if this pair of strategies is a SPE, it is worth making sure that they are clearly defined. We have to define the players' actions for *any* history. We

have done this as follows: in the first stage they both intend to cooperate. Then, at any stage they look back at history and ask: was there *ever* a deviation in *any previous* stage from always playing *mum*? If the answer is no, meaning that both players always cooperated and played *mum*, then in the next stage each player chooses *mum*.

If, however, there was *any deviation* from cooperation in the past, be it once or more often, be it by one player or both, the players will revert to playing *fink*, and by definition, stick to *fink* thereafter. For this reason, each of these strategies is commonly referred to as “grim trigger” strategies. They include a natural trigger: once someone deviates from *mum*, this is the trigger that causes the players to revert their behavior to *fink* forever. It is quite clear why the trigger is *grim*!

To verify that the grim trigger strategy pair is a SPE we need to check that there is no profitable deviation in any subgame! This task, however, is not that tedious as you may have expected given the one-stage deviation principle we discussed earlier for multistage games. Indeed, we have the following proposition:

**Proposition 13** *On an infinitely repeated game  $G(\delta)$  a profile of strategies  $s = (s_1, \dots, s_n)$  is a subgame perfect equilibrium if and only if there is no player  $i$  and no single history  $h_{t-1}$  for which player  $i$  would gain from deviating from  $s_i(h_{t-1})$ .*

**proof:** The “only if” part follows from the definition of a SPE. Indeed, if there were such a deviation then  $s$  could not have been a SPE. We are left to show that if it is not a SPE then there must be a one-stage deviation. (TO BE COMPLETED)

The one-stage deviation principle implies that all we have to do to confirm that a profile of strategies is a SPE is to check that no player has a history from which he would want to unilaterally deviate. Applying this to our repeated Prisoner’s dilemma may not seem helpful because there are still an infinite number of histories. Nonetheless, the task is not a hard one because even though there are an infinite number of histories, they fall into one of two *relevant classes*: either there was no deviation and the players intend to play  $M$ , or there was some previous decision and the players intend to play  $F$ .

Consider first the latter class of histories that *are not* consecutive sequences of only  $(M, M)$ . Notice that these histories are *off the equilibrium path* since following the proposed strategies would result in  $(M, M)$  being played forever. Thus, in *any subgame that is off the equilibrium path*, the proposed strategies recommend that each player play  $F$  in this period, and at any subsequent period since a deviation from  $M$  has occurred in the past. Clearly, no player would want to choose  $M$  instead of  $F$  since given her belief that her opponent will play  $F$  this will cause her a loss of  $-2$  (getting  $-1$  instead of  $1$ ) in this stage, and no gains in subsequent stages. Thus, in any subgame that is off the equilibrium path no player would ever want to unilaterally deviate.

Now consider the other class of histories that are consecutive sequences of  $(M, M)$ , which are all the histories that are *on the equilibrium path*. If a player chooses to play  $M$ , his current payoff is  $4$ , and his *continuation payoff* from the pair of strategies is an infinite sequence of  $4$ 's starting in the next period. Therefore his utility from following the strategy and not deviating will be

$$\begin{aligned} u_i^* &= 4 + \delta 4 + \delta^2 4 + \dots \\ &= 4 + \delta \frac{4}{1 - \delta} . \end{aligned}$$

The distinction of the player's payoff into today's payoff ( $4$ ) and the continuation payoff from following his strategy ( $\delta \frac{4}{1 - \delta}$ ) is useful. It recognizes the idea that a player's actions have two effects: The first is on his immediate payoffs today. The second is the effect on his *continuation equilibrium payoff*, which results from the way in which the equilibrium continues to unfold.

If the player deviates from  $M$  and chooses  $F$  instead then he gets  $5$  in the immediate stage, followed by his *continuation payoff* that is an infinite sequence of  $1$ 's. Indeed, we had already established that in the *continuation game* following a deviation, every player will stick to playing  $F$  since his opponent is expected to do so forever. Therefore his utility from deviating from the proposed strategy will be

$$\begin{aligned} u_i' &= 5 + \delta 1 + \delta^2 1 + \dots \\ &= 5 + \delta \frac{1}{1 - \delta} . \end{aligned}$$

Thus, we see what the trade-off boils down to: a deviation will yield an immediate gain of 1 since the player gets 5 instead of 4. However, the continuation payoff will then drop from  $\delta \frac{4}{1-\delta}$  to  $\delta \frac{1}{1-\delta}$ . We conclude that he will not want to deviate if  $u_i^* > u_i'$ , or

$$4 + \delta \frac{4}{1-\delta} \geq 5 + \delta \frac{1}{1-\delta} ,$$

which is equivalent to

$$\delta \geq \frac{1}{4} .$$

Once again, we see the value of *patience*. If the players are sufficiently patient, so that the future carries a fair amount of weight in their preferences, then there is a “reward and punishment” strategy that will allow them to cooperate *forever*. The loss from continuation payoffs will more than offset the gains from immediate defection, and this will keep the players on the cooperative path of play.

Using the probabilistic interpretation of  $\delta$ , these strategies will credibly allow the players to cooperate as long as the game is likely enough to be played again. In a way it is reminiscent of the old saying “what comes around goes around” to the extent that if the players are likely to meet again, good behavior can be rewarded and bad behavior can be punished. If the future does not offer that possibility, there can be no credible commitment to cooperation.

To relate this to the section on two-period multistage games, think of the repeated game as having two stages: “today”, which is the current stage, and “tomorrow”, which is the infinite sequence of play after this stage is over. For such multistage games we realized that there are two necessary conditions that follow if we could support non-Nash behavior in the first stage: multiple equilibria in the second stage, and sufficient patience on behalf of the players. Clearly, the inequality  $\delta \geq 0.25$  is the “sufficient patience” condition. But where are the multiple equilibria coming from? This is where the infinite repetition creates “magic”: from the unique equilibrium of the stage game we get multiple equilibria of the repeated game. This so called magic is so strong, that many types of behavior can be supported in a SPE as the next chapter demonstrates.

### 15.3 The Folk Theorem: Anything Goes<sup>♠</sup>

In this subsection we describe, in a somewhat formal way, one of the key insights about infinitely repeated games. This insight goes to the heart of the “magic” that when a stage game with a unique Nash equilibrium is repeated infinitely often, multiple SPE arise. To do this we introduce some more mathematical concepts that are needed for this section.

**Definition 33** Consider two vectors  $u = (u_1, u_2, \dots, u_n)$  and  $u' = (u'_1, u'_2, \dots, u'_n)$  in  $\mathbb{R}^n$ . We say that the vector  $\tilde{u} = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n)$  is a **convex combination** of  $u$  and  $u'$  if there exists some number  $\alpha \in [0, 1]$  such that  $\tilde{u} = \alpha u + (1 - \alpha)u'$ , or,  $\tilde{u}_i = \alpha u_i + (1 - \alpha)u'_i$  for all  $i \in \{1, \dots, n\}$ .

To think of this intuitively, consider two payoff vectors in the prisoners dilemma, say,  $(4, 4)$  and  $(5, -1)$ . If a line is drawn to connect between these in the plane  $\mathbb{R}^2$ , then any point on this line will be a convex combination of these two payoff vectors. For example, if we take  $\alpha = 0.3$  then the vector  $0.3(4, 4) + 0.7(5, -1) = (4.7, 0.5)$  is on that line. Another way to think of a convex combination is as a weighted average, where each point gets a (potentially) different weight. In this example,  $(4.7, 0.5)$  puts a weight of 0.3 on the vector  $(4, 4)$  and a weight of 0.7 on the vector  $(5, -1)$ .

**Definition 34** Given a set of vectors  $Z = \{u^1, u^2, \dots, u^k\}$  in  $\mathbb{R}^n$ , the **convex hull** of  $Z$  is,

$$\text{CoHull}(Z) = \left\{ u \in \mathbb{R}^n : \exists (\alpha_1, \dots, \alpha_k) \in \mathbb{R}_+^k, \sum_{j=1}^k \alpha_j = 1 \text{ such that } u = \sum_{j=1}^k \alpha_j u^j \right\}$$

We will apply this definition to the set of payoffs in a game. In a stage game with  $n$  players and finite strategy sets, the set of payoff vectors will be a finite set of vectors in  $\mathbb{R}^n$ , playing the role of  $Z$  in the definition. Then, we can think of all the convex combinations between all these payoff vectors. These together will define the convex hull of  $Z$ . For example, in the Prisoner’s dilemma given by the matrix

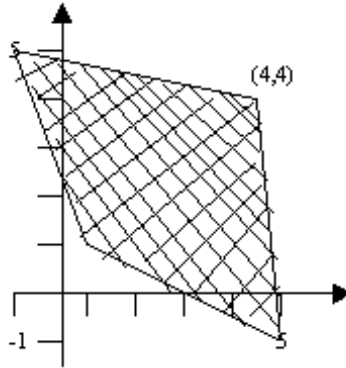


FIGURE 15.1.

	$M$	$F$
$M$	4, 4	-1, 5
$F$	5, -1	1, 1

the set of payoff vectors is  $Z = \{(4, 4), (5, -1), (-1, 5), (1, 1)\}$ , and the convex hull is given by the shaded area in figure 15.1.

At this stage you should suspect that there is a relationship between this convex hull and the set of *average payoffs* in the infinitely repeated game. Indeed, for a high enough discount factor *we can obtain any vector in the convex hull as an average payoff in the repeated game*. To prove this is not immediate, but the idea behind this fact is rather intuitive. Since any point in convex hull of payoffs is a weighted average of the stage game payoffs, we can achieve the appropriate weighting by repeating some sequence of the games payoffs that will imitate the required weights.

For this reason we call the convex hull of a game's payoffs the *feasible payoffs*, since they can be achieved in the infinitely repeated game. It turns out that many feasible payoffs can be obtained as a SPE of the repeated game. Legend says that the intuition for this result was floating around and is impossible to trace to a single contributor, hence it is referred to as the *Folk Theorem*. However, one of the earliest publications of such a result was written by James Friedman (1971), and we state a result in that spirit:

**Theorem 14 (The Folk Theorem)** *Let  $G$  be a finite, simultaneous move game of complete information, let  $(u_1^*, \dots, u_n^*)$  denote the payoffs from a Nash equilibrium of  $G$ , and let  $(u_1, \dots, u_n)$  be a feasible payoff of  $G$ . If  $u_i > e_i, \forall i \in N$  and if  $\delta$  is sufficiently close to 1, then there exists a SPE of the infinitely repeated game  $G(\delta)$  that achieves an average payoff arbitrarily close to  $(u_1, \dots, u_n)$ .*

L11

As with other important theorems, we will refrain from proving it but provide some intuition for it. Recall the idea behind SPE in multistage games: if we can find *credible* “reward and punishment” strategies then we can support behavior in an early stage that is not a Nash equilibrium of that stage. What we are doing here is choosing a path of play that yields strictly higher payoffs than a Nash equilibrium, and supporting it with the credible threat of reverting to the Nash equilibrium that is worse for all players. Since the game never ends, this threat does not hit an “unravelling” boundary that is a consequences of a last period.

An example is called for here. Recall the Prisoner’s dilemma above, and consider the average payoffs of (2,2). We will show that for a discount factor  $\delta$  that is close enough to 1, we can get average payoffs  $(\bar{u}_1, \bar{u}_2)$  that are arbitrarily close to (2, 2).

Take the following strategies:

- Player 1:
  - in period 1 play  $F$
  - in every even period play  $M$  if and only if the pattern of play was  $(F, M), (M, F), (F, M) \dots$ ; otherwise play  $F$
  - in every odd period after period 1 play  $F$
- Player 2:
  - in period 1 play  $M$
  - in every even period play  $F$
  - in every odd period after period 1, play  $M$  if and only if the pattern was  $(F, M), (M, F), (F, M) \dots$ ; otherwise play  $F$



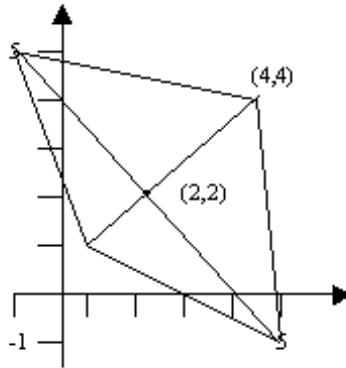


FIGURE 15.2.

As figure 15.2 demonstrates, these strategies will cause the payoff realizations to alternate between  $(5, -1)$  and  $(-1, 5)$ , resulting in the following average payoffs for the two players:

$$\begin{aligned}
 \bar{u}_1 &= (1 - \delta) \sum_{t=1}^{\infty} u_t^1 \\
 &= (1 - \delta)[5 + \delta(-1) + \delta^2(5) + \delta^3(-1) + \dots] \\
 &= (1 - \delta)[5(1 + \delta^2 + \delta^4 + \dots) + (-1)(\delta + \delta^3 + \delta^5 + \dots)] \\
 &= (1 - \delta) \left[ 5 \cdot \frac{1}{1 - \delta^2} + (-1) \cdot \frac{\delta}{1 - \delta^2} \right] \\
 &= \frac{5 - \delta}{1 + \delta},
 \end{aligned}$$

and similarly for player 2,

$$\begin{aligned}
 \bar{u}_2 &= (1 - \delta)[(-1) + \delta(5) + \delta^2(-1) + \delta^3(5) + \dots] \\
 &= \frac{-1 + 5\delta}{1 + \delta}.
 \end{aligned}$$

Now notice that

$$\lim_{\delta \rightarrow 1} \bar{u}_1 = 2 = \lim_{\delta \rightarrow 1} \bar{u}_2$$

so that average payoffs arbitrarily close to (2,2) can be obtained for  $\delta$  close enough to 1.<sup>4</sup>

Are these strategies SPE? As before, we need to check that no player would want to unilaterally deviate at any stage of the game, and the fact that every subgame is the same infinitely repeated game is very helpful. It suffices to check for 3 types of deviations:

1. If there was a deviation and we enter the punishment phase ( $F$  forever), no player will want to deviate since this guarantees a loss of  $-2$  in the deviation period with no gains (like in the example we did earlier).
2. Does player 1 (player 2) want to deviate in an odd (even) period? If she does not deviate then she obtains an average payoff of  $\bar{u}_1$ . If she deviates, she will get 4 instead of 5 in the deviating period, but in subsequent periods the players will play ( $F, F$ ) forever, giving a stream of 1's in each period instead of alternating between  $(-1)$  and 5, so this deviation yields an average payoff of

$$u'_1 = (1 - \delta)[4 + \delta \cdot 1 + \delta^2 \cdot 1 + \dots] = (1 - \delta)4 + \delta .$$

It is easy to see that for high  $\delta$  close to 1 the value of  $u'_1$  will be itself just above 1, and  $\bar{u}_1$  is close to 2, implying that this deviation is not profitable. More precisely, the player will not want to deviate if and only if  $\frac{5-\delta}{1+\delta} \geq (1-\delta)4 + \delta$ , which holds for all  $\delta \in [0, 1]$ .

3. Does Player 1 (player 2) want to deviate in any even (odd) period? If she does not deviate then she obtains an average payoff of  $\frac{-1+5\delta}{1+\delta}$ . If she deviates, she will get 1 instead of  $(-1)$  in the deviating period, but in subsequent periods the players will play ( $F, F$ ) forever, giving a stream of 1's in each

---

<sup>4</sup>(Technical point:) In fact, we can play with the strategies to get the average payoff closer to (2,2) without increasing  $\delta$ . To see this, imagine that  $\delta = 0.99$  so that the strategies above give us average payoffs of  $\bar{u}_1 = 2.0151$  and  $\bar{u}_2 = 1.9849$ . Now imagine that we keep these strategies except for one modification: in period 595, which is an odd period, have player 1 play  $M$  instead of  $F$  and have player 2 play  $F$  instead of  $M$ . This means that in period 595 player 1 gets  $(-1)$  instead of 5, and the opposite is true for player 2. This means that player 1's (2's) average payoff decreases (increases) by  $(0.99)^{595} \times 6 = 0.015174$ . Thus, the average payoffs would be  $\bar{u}_1 = 1.9999$  and  $\bar{u}_2 = 2.0001$ . We can continue fine tuning the payoffs to actually achieve (2,2)!

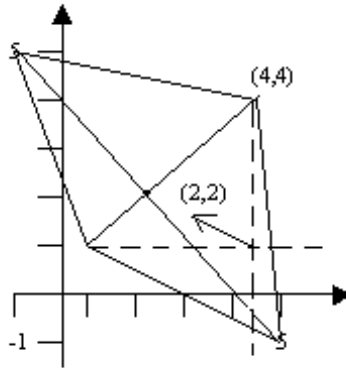


FIGURE 15.3.

period instead of alternating between  $(-1)$  and  $5$ , so this deviation yields an average payoff of

$$u_1'' = (1 - \delta)[1 + \delta \cdot 1 + \delta^2 \cdot 1 + \dots] = 1 .$$

Thus, she will not deviate if  $\frac{-1+5\delta}{1+\delta} \geq 1$ , which holds if and only if  $\delta \geq \frac{1}{2}$ .

Thus, we have concluded that the payoffs of  $(2, 2)$  can (nearly) be achieved by the pair of strategies that we have specified when the discount factor is high enough. The infinite repetition introduced the ability to include “reward and punishment” strategies, which are intuitive incentive schemes to use when people interact often, and believe that every interaction is likely to be followed by another one.

### 15.4 Additional Material:

- Repeated game theory essentially explains the whole cross-hatched area, based on beliefs. There are a lot of applications for this, e.g. explaining transition economies versus developed economics, both at equilibriums but where one is at a higher state than the other.
- Interpret different equilibria with different beliefs, i.e., Norms (Greif)

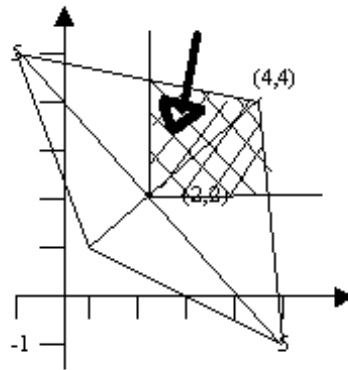


FIGURE 15.4.