Infinitely Repeated Games with Discounting

Introduction

We'll now discuss repeated games which are "infinitely repeated." This need not mean that the game never ends, however. We will see that this framework is appropriate for modeling situations in which the game eventually ends (with probability one) but the players are uncertain about exactly when the last period is (and they always believe there's some chance the game will continue to the next period).

We'll call the stage game G and interpret it to be a simultaneous-move matrix game which remains exactly the same through time. As usual we let the player set be $I=\{1,\ldots,n\}$. Each player has a pure action space A_i.[1] The space of action profiles is $A=\times_{i\in I}A_i$. Each player has a von Neumann-Morgenstern utility function defined over the outcomes of G, $g_i\colon A\rightarrow\mathbb{R}$.

The stage game repeats each period, starting at $t=0$. Although each stage game is a simultaneous-move game, so that each player acts in ignorance of what her opponent is doing that period, we make the "observable action" or "standard signaling" assumption that the play which occurs in each repetition of the stage game is revealed to all the players before the next repetition. Combined with perfect recall, this allows a player to condition her current action on all earlier actions of her opponents.

We can think of the players as receiving their stage-game payoffs period-by-period. Their repeated-game payoffs will be an additively separable function of these stage-game payoffs. Right away we see a potential problem: There is an infinite number of periods and, hence, of stage-game payoffs to be added

1 I want to reserve "s" and "S" to represent typical strategies and strategy spaces, respectively, in the repeated game. So I'm using "a" and "A" for stage-game actions and action spaces.

up. In order that the players' repeated-game payoffs be well defined we must ensure that this infinite sum does not blow up to infinity. We ensure the finiteness of the repeated-game payoffs by introducing discounting of future payoffs relative to earlier payoffs. Such discounting can be an expression of time preference and/or uncertainty about the length of the game. We introduce the average discounted payoff as a convenience which normalizes the repeated-game payoffs to be "on the same scale" as the stage-game payoffs.

Infinite repetition can be key to obtaining behavior in the stage games which could not be equilibrium behavior if the game were played once or a known finite number of times. For example, finking every period by both players is the unique equilibrium in any finite repetition of the prisoners' dilemma.[2] When repeated an infinite number of times, however, cooperation in every period *is* an equilibrium if the players are "sufficiently patient."

We first show that cooperation is a Nash equilibrium of the infinitely repeated game for sufficiently patient players. Then we discuss subgames of infinitely repeated games and show that, in our uniformly discounted payoff scenario, a subgame is the same infinitely repeated game as the original game. We derive how to restrict a repeated-game strategy to a subgame. This allows us to complete the analysis of the example by showing that cooperation in every period by both players is a subgame-perfect equilibrium of the infinitely repeated prisoners' dilemma for sufficiently patient players.

Discounting the future

When we study infinitely repeated games we are concerned about a player who receives a payoff in each of infinitely many periods. In order to represent her preferences over various infinite payoff streams we want to meaningfully summarize the desirability of such a sequence of payoffs by a single number. A common assumption is that the player wants to maximize a weighted sum of her per-period payoffs, where she weights later periods less than earlier periods.

For simplicity this assumption often takes the particular form that the sequence of weights forms a geometric progression: for some fixed $\delta \in (0,1)$, each weighting factor is δ times the previous weight. (δ is called her *discount factor*.) If in each period t player i receives the payoff u_i^t, we could summarize the desirability of the payoff stream $u_i^0, u_i^1, \ldots$ by the number[3,4]

$$\sum_{t=0}^{\infty} \delta^t u_i^t. \tag{1}$$

Such an intertemporal preference structure has the desirable property that the infinite sum of the

---

[2] See Theorem 4 of "Repeated Games."

[3] It's imperative that you not confuse superscripts with exponents! You should be able to tell from context. For example, in the summation (1), the t in "δ^t" is an exponent; in x_i^t, it is a superscript identifying the time period.

[4] Note that this assigns a weight of unity to period zero. It's the relative weights which are important (why?), so we are free to normalize the weights any way we want.

weighted payoffs will be finite (since the stage-game payoffs are bounded). You would be indifferent between a payoff of x^t at time t and a payoff of $x^{t+\tau}$ received τ periods later if

$$x^t = \delta^\tau x^{t+\tau}. \tag{2}$$

A useful formula for computing the finite and infinite discounted sums we will encounter is

$$\sum_{t=T_1}^{T_2} \delta^t = \frac{\delta^{T_1} - \delta^{T_2+1}}{1-\delta}, \tag{3}$$

which, in particular, is valid for $T_2 = \infty$.[5]

Interpreting the discount factor

One way we can interpret the discount factor δ is as an expression of traditional time preference. For example, if you received a dollar today you could deposit it in the bank and it would be worth $\$(1+r)$ tomorrow, where r is the per-period rate of interest. You would be indifferent between a payment of $\$x^t$ today and $\$x^{t+\tau}$ received τ periods later only if the future values were equal: $(1+r)^\tau x^t = x^{t+\tau}$. Comparing this indifference condition with (2), we see that the two representations of intertemporal preference are equivalent when

$$\delta = \frac{1}{1+r}. \tag{4}$$

We can relate a player's discount factor δ to her *patience*. How much more does she value a dollar today, at time t, than a dollar received $\tau > 0$ periods later? The relative value of the later dollar to the earlier is $\delta^{t+\tau}/\delta^t = \delta^\tau$. As $\delta \to 1$, $\delta^\tau \to 1$, so as her discount factor increases she values the later amount more and more nearly as much as the earlier payment. A person is more patient the less she minds waiting for something valuable rather than receiving it immediately. So we interpret higher discount factors as higher levels of patience.

There's another reason you might discount the future. You may be unsure about how long the game will continue. Even in the absence of time preference *per se*, you would prefer a dollar today rather than a promise of one tomorrow because you're not sure tomorrow will really come. (A payoff at a future time is really a conditional payoff—conditional on the game lasting that long.)

For example, assume that your beliefs about the length of the game can be represented by a constant *conditional continuation probability* ρ; i.e. conditional on time t being reached, the probability that the game will continue at least until the next period is ρ.[6] Without loss of generality we assign a probability

5 See the "Appendix: Discounting Payoffs" for a derivation of this formula.

6 Note that there is no upper bound which limits the length of the game; for any value of $T \in \mathbb{Z}_{++}$, there is a positive probability of the game lasting more than T periods. If, on the other hand, players' beliefs about the duration are such that there is a common knowledge upper bound $\overline{T}$ on the length of the game, the game is similar to a finitely repeated game with a common knowledge and definite

of one to the event that we reach period zero.[7] Denote by $\mathbb{P}(t)$ the probability that period t is reached and by $\mathbb{P}(t\,|\,t')$ the probability that period t is reached conditional on period t' being reached; so $\mathbb{P}(t+1\,|\,t)=\rho$ for all t. The probability of reaching period 1 is $\mathbb{P}(1)=\mathbb{P}(1\,|\,0)\mathbb{P}(0)=\rho\cdot 1=\rho$. The probability of reaching period 2 is $\mathbb{P}(2)=\mathbb{P}(2\,|\,1)\mathbb{P}(1)=\rho^2$. In general, the probability of reaching period t is $\mathbb{P}(t)=\rho^t$. The probability that the game ends is $1-\mathbb{P}(\infty)=1$, i.e. one minus the probability that it continues forever; therefore the game is finitely repeated with probability one.

Consider the prospect of receiving the conditional payoff of one dollar at time t, given your time preference δ and the conditional continuation probability ρ. The probability that the payoff will be received is the probability that period t is reached, viz. ρ^t. Its discounted value if received is δ^t and zero if not received. Therefore its expected value is $\rho^t\delta^t=\tilde{\delta}^t$, where $\tilde{\delta}\equiv\rho\delta$ is the discount factor which is equivalent to the combination of the time preference with continuation uncertainty. So, even though the game is finitely repeated, due to the uncertainty about its duration we can model it as an infinitely repeated game.

The average discounted payoff

If we adopted the summation (1) as our players' repeated-game utility function, and if a player received the same stage-game payoff v_i in every period, her discounted repeated-game payoff, using (3), would be $v_i/(1-\delta)$. This would be tolerable, but we'll find it more convenient to transform the repeated-game payoffs to be "on the same scale" as the stage-game payoffs, by multiplying the discounted payoff sum from (1) by $(1-\delta)$. So we define the *average discounted value* of the payoff stream $u_i^0, u_i^1,\ldots$ by

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t u_i^t. \tag{5}$$

Now if a player receives the stage-game payoff v_i in every period, her repeated-game payoff is v_i as well. What could be simpler?[8]

Let's make some sense of why (5) is interpreted as an average payoff. Consider some time varying sequence of payoffs u_t, for $t=0,1,2,\ldots$, and let V be the resulting value of the payoff function (5). Now replace every per-period payoff u_t with the value V. The value of the payoff function (5) associated with this new constant-payoff stream of $V,V,V,\ldots$ is

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t V=V(1-\delta)\sum_{t=0}^{\infty}\delta^t=V.$$

---

duration in the following sense: If the stage game has a unique Nash equilibrium payoff vector, every subgame-perfect equilibrium of the repeated game involves Nash equilibrium stage-game behavior in every period.

7 If we don't get to period zero, we never implement our strategies. Therefore we can assume that we have reached period zero if we do indeed implement our strategies.

8 We will see at a later date that an additional advantage to this normalization is that the convex hull of the feasible repeated-game payoffs using the average discounted value formulation is invariant to the discount factor.

In other words, the average discounted payoff of a payoff stream $u_0, u_1, u_2, \ldots$ is that value V which, if you received it every period, would give you the same average discounted payoff as the original stream.

We'll often find it convenient to compute the average discounted value of an infinite payoff stream in terms of a leading finite sum and the sum of a trailing infinite substream. For example, say that the payoffs v_i^t a player receives are some constant payoff v_i' for the first t periods, viz. $0, 1, 2, \ldots, t-1$, and thereafter she receives a different constant payoff v_i'' in each period $t, t+1, t+2, \ldots$. The average discounted value of this payoff stream is

$$(1-\delta)\sum_{\tau=0}^{\infty} \delta^\tau v_i^\tau = (1-\delta)\left(\sum_{\tau=0}^{t-1} \delta^\tau v_i^\tau + \sum_{\tau=t}^{\infty} \delta^\tau v_i^\tau\right) = (1-\delta)\left(\frac{v_i'(1-\delta^t)}{1-\delta} + \frac{v_i''\delta^t}{1-\delta}\right) = (1-\delta^t)\,v_i' + \delta^t v_i''. \tag{6}$$

We see that the average discounted value of this stream of bivalued stage-game payoffs is a convex combination of the two stage-game payoffs.

We can iterate this procedure in order to evaluate the average discounted value of more complicated payoff streams. For example—and this one will be particularly useful—say you receive v_i' for the t periods $0, 1, \ldots, t-1$ as before, and then you receive v_i'' only in period t and receive v_i''' every period thereafter. The average discounted value of the stream beginning in period t (discounted to period t) is $(1-\delta)v_i'' + \delta v_i'''$. Substituting this for v_i'' in (6), we find that the average discounted value of this three-valued payoff stream is

$$(1-\delta^t)v_i' + \delta^t[(1-\delta)v_i'' + \delta v_i''']. \tag{7}$$

Example: Cooperation in the Prisoners' Dilemma (Nash Equilibrium)

In the one-shot prisoners' dilemma the players cannot avoid choosing their dominant strategy Fink. (See Figure 2.) Even when this game is finitely repeated—because the stage game has a unique Nash equilibrium—the unique subgame-perfect equilibrium has both players Finking every period.[9] However, when the players are sufficiently patient we can sustain cooperation (i.e. keeping Mum) in every period as a subgame-perfect equilibrium of the infinitely repeated game. First we will see that such cooperation is a Nash equilibrium of the repeated game. Later we will show that this cooperation is a subgame-perfect equilibrium.

	M	*F*
M	1, 1	−1, 2
F	2, −1	0, 0

Figure 2: A Prisoners' Dilemma

When an infinitely repeated game is played, each player i has a repeated-game strategy s_i, which is a sequence of history-dependent stage-game strategies s_i^t; i.e. $s_i = (s_i^0, s_i^1, \ldots)$, where each $s_i^t\colon A^t \to A_i$. The

9 See Theorem 4 in the "Repeated Games."

n-tuple of individual repeated-game strategies is the repeated-game strategy profile s; i.e. $s=(s_1,\ldots,s_n)$.

The repeated-game strategies I exhibit, which are sufficient to sustain cooperation, have the following form: Cooperate (i.e. play Mum) in the first period. In later periods, cooperate if both players have always cooperated. However, if either player has ever Finked, Fink for the remainder of the game. More precisely and formally, we write player i's repeated-game strategy $\bar{s}_i=(\bar{s}_i{}^0,\bar{s}_i{}^1,\ldots)$ as the sequence of history-dependent stage-game strategies such that in period t and after history h^t,

$$\bar{s}_i^t(h^t)=\begin{cases} M, & t=0 \quad \text{or} \quad h^t=\left((M,M)^t\right), \\ F, & \text{otherwise.} \end{cases} \tag{8}$$

Recall that the history h^t is the sequence of stage-game action profiles which were played in the t periods $0,1,2,\ldots,t-1$. "$((M,M)^t)$" is just a way of writing "$((M,M),(M,M),\ldots,(M,M))$", where "$(M,M)$" is repeated t times.

We can simplify (8) further in a way that will be useful in our later analysis: Because h^0 is the null history, we adopt the convention that, for any action profile $a\in A$, the statement $h^0=(a)^0$ is true.[10] With this understanding, $t=0$ implies that $h^t=((M,M)^t)$. Therefore

$$\bar{s}_i^t(h^t)=\begin{cases} M, & h^t=\left((M,M)^t\right), \\ F, & \text{otherwise.} \end{cases} \tag{9}$$

First, I'll show that for sufficiently patient players the strategy profile $\bar{s}=(\bar{s}_1,\bar{s}_2)$ is a Nash equilibrium of the repeated game. Later I'll show that for the same required level of patience these strategies are also a subgame-perfect equilibrium.

If both players conform to the alleged equilibrium prescription, they both play Mum at $t=0$. Therefore at $t=1$, the history is $h^1=(M,M)$; so they both play Mum again. Therefore at $t=2$, the history is $h^2=((M,M),(M,M))$, so they both play Mum again. And so on…. So the path of s is the infinite sequence of cooperative action profiles $((M,M),(M,M),\ldots)$. The repeated-game payoff to each player corresponding to this path is trivial to calculate: They each receive a payoff of 1 in each period, therefore the average discounted value of each player's payoff stream is 1.

Can player i gain from deviating from the repeated-game strategy $\bar{s}_i$ given that player j is faithfully following $\bar{s}_j$? Let t be the period in which player i first deviates. She receives a payoff of 1 in the first t periods $0,1,\ldots,t-1$. In period t, she plays Fink while her conforming opponent played Mum, yielding player i a payoff of 2 in that period. This defection by player i now triggers an open-loop Fink-always response from player j. Player i's best response to this open-loop strategy is to Fink in every period

10 Here is a case of potential ambiguity between time superscripts and exponents. "$(a)^0$" means "a" raised to the zero power; i.e. the profile a is played zero times.

herself. Thus she receives zero in every period $t+1, t+2, \ldots$. To calculate the average discounted value of this payoff stream to player i we can refer to (7), and substitute $v_i{'}=1$, $v_i{''}=2$, and $v_i{'''}=0$. This yields player i's repeated-game payoff when she defects in period t in the most advantageous way to be $1-\delta^t(2\delta-1)$. This is weakly less than the equilibrium payoff of 1, for any choice of defection period t, as long as $\delta \geq \frac{1}{2}$. Thus we have defined what I meant by "sufficiently patient:" cooperation in this prisoners' dilemma is a Nash equilibrium of the repeated game as long as $\delta \geq \frac{1}{2}$.

Restricting strategies to subgames

Consider the subgame which begins in period τ after a history $h^\tau=(a^0, a^1, \ldots, a^{\tau-1})$. This subgame is itself an infinitely repeated game; in fact, it is the same infinitely repeated game as the original game in the sense that 1 the same stage game G is played relentlessly and 2 a unit payoff tomorrow is worth the same as a payoff of δ received today. Another way to appreciate 2 is to consider an infinite stream of action profiles $a^0, a^1, \ldots$. If a player began in period τ to receive the payoffs from this stream of action profiles, the contribution to her total repeated-game payoff would be δ^τ times its contribution if it were received beginning at period zero. If player i were picking a sequence of history-dependent stage-game strategies to play beginning when some subgame h^τ were reached, against fixed choices by her opponents, she would face exactly the same maximization problem as she would if she thought she was just starting the game.[11]

What behavior in the subgame is implied by the repeated-game strategy profile s? The restriction of the repeated-game strategy profile s to the subgame h^τ tells us how the subgame will be played. Let's denote this restriction by $\hat{s}$. If we played the infinitely-repeated game from the beginning by having each player play her part of $\hat{s}$, we would see the same infinite sequence of action profiles as we would see if instead s were being played and we began our observation at period τ.

For example, in the initial period of the subgame (viz. period τ on the original game's timeline and period zero on the subgame's timeline), player i chooses the action dictated by her subgame strategy $\hat{s}_i$, viz. $\hat{s}_i{}^0$. This action must be what her original repeated-game strategy s_i would tell her to do after the history h^τ by which this subgame was reached. I.e.

$$\hat{s}_i{}^0 = s_i{}^\tau(h^\tau). \tag{10}$$

After all the players choose their actions, a stage-game action profile results. We can denote this action profile in two ways: as $\hat{a}^0$, which refers to the initial period of the subgame, or as a^τ, which refers to the timeline of the original game. (See Figure 3.) This action profile $\hat{a}^0=a^\tau$ then becomes part of the history which conditions the next period's choices. Period 1 of the subgame is period $\tau+1$ of the original game.

---

[11] This paragraph may be unclear, but it's not a trivial statement! Its validity relies crucially on the structure of the weighting coefficients in the infinite summation. Let α_t be the weight for period t's per-period payoff. In the discounted payoff case we consider here, we have $\alpha_t=\delta^t$. The fact that the "continuation game" is the same game as the whole game is a consequence of the fact that, for all t and τ, the ratio $\alpha_{t+\tau}/\alpha_t$ depends only upon τ but not upon t. In our case, this ratio is just δ^τ. On the other hand, for example, if the weighting coefficients were $\alpha_t=\delta^t/t$, we could not solve each continuation game just as another instance of the original game.

The history in the subgame is $\hat{h}^1=(\hat{a}^0)$; the history in the original game is

$$h^{\tau+1}=(a^0,a^1,\ldots,a^{\tau-1},a^\tau)=(h^\tau;\ a^\tau)=(h^\tau;\ \hat{a}^0)=(h^\tau;\ \hat{h}^1), \tag{11}$$

where $(h^\tau;\ a^\tau)$ is the *concatenation* of h^τ with a^τ.[12] The action chosen next by player i can be denoted equivalently as $\hat{s}_i^{\,1}(\hat{h}^1)$ or $s_i^{\,\tau+1}((h^\tau;\ \hat{h}^1))$. Therefore we must have

$$\hat{s}_i^{\,1}(\hat{h}^1)=s_i^{\,\tau+1}((h^\tau;\ \hat{h}^1)). \tag{12}$$

Continuing the analysis period-by-period, we see that we can define the restriction of s_i to the subgame determined by h^τ by

$$\hat{s}_i^{\,t}(\hat{h}^t)=s_i^{\,\tau+t}((h^\tau;\ \hat{h}^t)), \tag{13}$$

for every $t\in\{0,1,\ldots\}$.

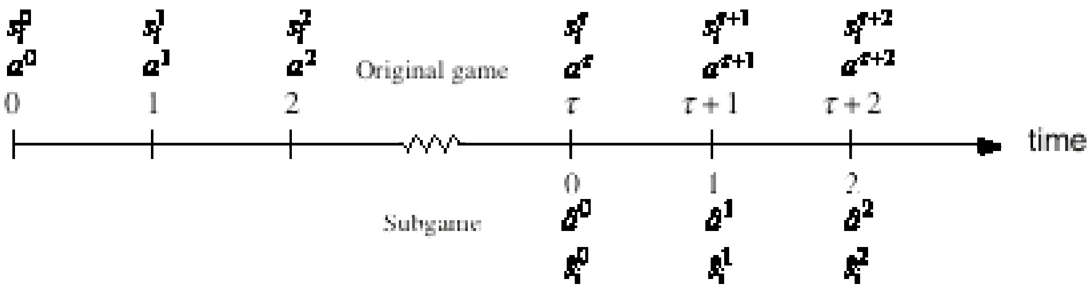

Figure 3: Histories and history-dependent stage-game strategies in two time frames: the original game's and a subgame's.

Example: Cooperation in the Prisoners' Dilemma (Subgame-perfect Equilibrium)

To verify that $\bar{s}$ is a subgame-perfect equilibrium of the repeated prisoners' dilemma we need to check that this strategy profile's restriction to each subgame is a Nash equilibrium of that subgame. Consider a subgame, beginning in period τ with some history h^τ. What is the restriction of $\bar{s}_i$ to this subgame? Denoting the restriction by $\hat{s}_i$ and using definitions (13) and (9), we find

$$\hat{s}_i^t(\hat{h}^t)=\bar{s}_i^{t+\tau}((h^\tau;\ \hat{h}^t))=\begin{cases} M, & h^\tau=\big((M,M)^\tau\big) \quad \text{and} \quad \hat{h}^t=\big((M,M)^t\big), \\ F, & \text{otherwise.}\end{cases} \tag{14}$$

We can partition the subgames of this game, each identified by a beginning period τ and a history h^τ, into two classes: A those in which both players chose Mum in all previous periods, i.e. $h^\tau=((M,M)^\tau)$, and B those in which a defection by either player has previously occurred.[13]

For those subgames in class A, the sequence of restrictions $\hat{s}_i^{\,t}(\hat{h}^t)$ from (14) reduces to the sequence of

[12] To *concatenate* two sequences a with b is simply to append b to the end of a.

[13] Note that class A includes the "subgame of the whole," which begins in period zero. (Because there are no previous periods, it is trivially true that (M,M) was played in all previous periods.)

original stage-game strategies $\bar{s}_i^{\,t}(h^t)$ from (9). I.e. for all τ and $h^\tau = ((M,M)^\tau)$,[14][15]

$$\hat{s}_i^t(\hat{h}^t) = \begin{cases} M, & h^\tau = \big((M,M)^\tau\big) \quad \text{and} \quad \hat{h}^t = \big((M,M)^t\big), \\ F, & \text{otherwise.} \end{cases} = \begin{cases} M, & \hat{h}^t = \big((M,M)^t\big), \\ F, & \text{otherwise.} \end{cases} = \bar{s}_i^t(\hat{h}^t) \qquad (15)$$

Because $\bar{s}$ is a Nash-equilibrium strategy profile of the repeated game, for each subgame h^τ in class A, the restriction $\hat{s}$ is a Nash-equilibrium strategy profile of the subgame when $\delta \geq \frac{1}{2}$.

For any subgame h^τ in class B, $h^\tau \neq ((M,M)^\tau)$. Therefore the restriction $\hat{s}_i$ of $\bar{s}_i$ specifies $\hat{s}_i^{\,t} = F$ for all $t \in \{0, 1, \ldots\}$. In other words, in any subgame reached by some player having Finked in the past, each player chooses the open-loop strategy "Fink always." Therefore the repeated-game strategy profile $\hat{s}$ played in such a subgame is an open-loop sequence of stage-game Nash equilibria. From Theorem 1 of the Repeated Games handout we know that this is a Nash equilibrium of the repeated game and hence of this subgame.

So we have shown that for every subgame the restriction of $\bar{s}$ to that subgame is a Nash equilibrium of that subgame for $\delta \geq \frac{1}{2}$. Therefore $\bar{s}$ is a subgame-perfect equilibrium of the infinitely repeated prisoners' dilemma when $\delta \geq \frac{1}{2}$.

14 When $h^\tau = ((M,M)^\tau)$, $[h^\tau = ((M,M)^\tau)$ and $h^t = ((M,M)^t)] \Leftrightarrow h^t = ((M,M)^t)$.

15 There is no distinction between the roles played between h^t and $\hat{h}^t$ in these two expressions. They are both dummy variables.

Appendix: Discounting Payoffs

Often in economics—for example when we study repeated games—we are concerned about a player who receives a payoff in each of many (perhaps infinitely many) periods and we want to meaningfully summarize her entire sequence of payoffs by a single number. A common assumption is that the player wants to maximize a weighted sum of her per-period payoffs, where she weights later periods less than earlier periods. For simplicity this assumption often takes the particular form that the sequence of weights forms a geometric progression: for some fixed $\delta \in (0,1)$, each weighting factor is δ times the previous weight. (δ is called her *discount factor*.) In this handout we're going to derive some useful results relevant to this kind of payoff structure. We'll see that you really don't need to memorize a bunch of different formulas. You can easily work them all out from scratch from a simple starting point.

Let's set this problem up more concretely. It will be convenient to call the first period $t=0$. Let the player's payoff in period t be denoted u_t. By convention we attach a weight of 1 to the period 0 payoff.[16] The weight for the next period's payoff, u_1, then, is $\delta \cdot 1 = \delta$; the weight for the next period's payoff, u_2, then, is $\delta \cdot \delta = \delta^2$; the weight for the next period's payoff, u_3, then, is $\delta \cdot \delta^2 = \delta^3$. You see the pattern: in general the weight associated with payoff u_t is δ^t. The weighted sum of all the payoffs from $t=0 \to t=T$ (where T may be ∞) is

$$u_0 + \delta u_1 + \delta^2 u_2 + \cdots + \delta^T u_T = \sum_{t=0}^{T} \delta^t u_t. \tag{1}$$

The infinite summation of the discount factors

We sometimes will be concerned with the case in which the player's period payoffs are constant over some stretch of time periods. In this case the common u_t value can be brought outside and in front of the summation over those periods because those u_t values don't depend on the summation index t. Then we will be essentially concerned with summations of the form

In a hurry?

Here's a formula we'll eventually derive and from which you can derive a host of other useful discounting expressions. In particular, it's valid when $T_2 = \infty$.

$$\sum_{t=T_1}^{T_2} \delta^t = \frac{\delta^{T_1} - \delta^{T_2+1}}{1-\delta}.$$

16 It's the relative weights which are important (why?), so we are free to normalize the sequence of weights any way we want.

$$\sum_{t=T_1}^{T_2} \delta^t. \tag{2}$$

To begin our analysis of these summations we'll first study a special case:

$$\sum_{t=0}^{\infty} \delta^t, \tag{3}$$

and then see how all of the other results we'd like to have can be easily derived from this special case.

The only trick is to see that the infinite sum (3) is mathemagically equal to the reciprocal of $(1-\delta)$, i.e.

$$\frac{1}{1-\delta} = 1 + \delta + \delta^2 + \cdots = \sum_{t=0}^{\infty} \delta^t. \tag{4}$$

To convince yourself that the result of this simple division really is this infinite sum you can use either of two heuristic methods. First, there's *synthetic division:*[17]

$$\begin{array}{r l}
 & 1+\delta+\delta^2+\cdots \\ \cline{2-2}
1-\delta \big) & 1 \\
 & \underline{1-\delta} \\
 & \quad \delta \\
 & \quad \underline{\delta-\delta^2} \\
 & \qquad \delta^2 \\
 & \qquad \underline{\delta^2-\delta^3} \\
 & \qquad\qquad \cdot
\end{array} \tag{5}$$

If you don't remember synthetic division from high school algebra, you can alternatively verify (4) by multiplication:[18]

$$\begin{aligned}
(1-\delta)(1+\delta+\delta^2+\cdots) &= (1+\delta+\delta^2+\cdots)-\delta(1+\delta+\delta^2+\cdots) \\
&= (1+\delta+\delta^2+\cdots)-(\delta+\delta^2+\cdots) \\
&= 1.
\end{aligned} \tag{6}$$

An infinite summation starting late

We'll encounter cases where a player's per-period payoffs vary in the early periods, say up through

17 I learned this in tenth grade, and I haven't seen it since; so I can't give a rigorous justification for it. In particular: its validity in this case must depend on $\delta < 1$, but I don't know a precise statement addressing this.

18 Here we rely on $\delta < 1$ to guarantee that the two infinite sums on the right-hand side converge and therefore that we can meaningfully subtract the second from the first.

period T, but after that she receives a constant payoff per period. To handle this latter constant-payoff stage we'll want to know the value of the summation

$$\sum_{t=T+1}^{\infty} \delta^t. \tag{7}$$

Now that we know the value from (4) for the infinite sum that starts at $t=0$, we can use that result and the technique of *change of variable* in order to find the value of (7).

If only the lower limit on the summation (7) were $t=0$ instead of $t=T+1$, we'd know the answer immediately. So we define a new index τ, which depends on t, and transform the summation (7) into another summation whose index τ runs from $0\to\infty$ and whose summand is δ^τ. So, let's define

$$\tau = t-(T+1). \tag{8}$$

We chose this definition because when $t=T+1$—like it does for the first term of (7)—then $\tau=0$, which is exactly what we want τ to be for the first term of the new summation.[19] For purposes of later substitution we write (8) in the form

$$t=\tau+(T+1). \tag{9}$$

Substituting for t using (9), summation (7) can now be rewritten as

$$\sum_{t=T+1}^{\infty} \delta^t = \sum_{\tau+(T+1)=T+1}^{\infty} \delta^{\tau+(T+1)} = \delta^{T+1} \sum_{\tau=0}^{\infty} \delta^\tau = \frac{\delta^{T+1}}{1-\delta}. \tag{10}$$

where we used (4) to transform the final infinite sum into its simple reciprocal equivalent.[20]

The finite summation

Our game might only last a finite number of periods. In this case we want to know the value of

$$\sum_{t=0}^{T} \delta^t. \tag{11}$$

We can immediately get this answer by combining the two results—(4) and (10)—which we've already found. We need only note that the infinite sum in (3) is the sum of two pieces: 1 the piece we want in (11) and 2 the infinite piece in (7). So we immediately obtain

$$\sum_{t=0}^{T} \delta^t = \sum_{t=0}^{\infty} \delta^t - \sum_{t=T+1}^{\infty} \delta^t = \frac{1}{1-\delta} - \frac{\delta^{T+1}}{1-\delta} = \frac{1-\delta^{T+1}}{1-\delta}. \tag{12}$$

---

19 Of course, we could have achieved this goal by defining $\tau=(T+1)-t$. But we want τ to *increase* with t, so that when $t\to\infty$, we'll also have $\tau\to\infty$.

20 The "t" in (4) is just a dummy variable; (4) is still true if you replace every occurrence of "t" with "τ".

Endless possibilities

The techniques we have used here of employing a change of variable and combining previous results is very general and can be used to obtain the value of summations of the general form in (2). The particular results we derived, viz. (10) and (12), are merely illustrative cases.

As an exercise, you should be able to easily combine results we have already found here to show that the value of the general summation in (2) is

$$\sum_{t=T_1}^{T_2} \delta^t = \frac{\delta^{T_1} - \delta^{T_2+1}}{1-\delta}. \tag{13}$$

Note that (13) is valid even when $T_2 = \infty$, because

$$\sum_{t=T_1}^{\infty} \delta^t = \lim_{T_2 \to \infty} \sum_{t=T_1}^{T_2} \delta^t = \lim_{T_2 \to \infty} \frac{\delta^{T_1} - \delta^{T_2+1}}{1-\delta} = \frac{\delta^{T_1}}{1-\delta}, \tag{14}$$

where we used the fact that

$$\lim_{t \to \infty} \delta^t = 0, \tag{15}$$

for $\delta \in [0, 1)$.

You should verify that, by assigning the appropriate values to T_1 and T_2 in (13), we can recover all of our previous results, viz. (4), (10), and (12).