

# Iterated Prisoner's Dilemma

Peter Varshavsky

Version of June 10, 2014

# Contents

## 0.1 Notes from June 3, 2014

### 0.1.1 What to ask

- Scope.
- Amount of proof and rigor.
- Audience.

### 0.1.2 FOCUS

- Exposition:
- Purpose: my reading of the history of the problem.
- Replicator dynamics: originally probably from Darwinian biological systems, but nothing prevents replicator dynamics to be combined with sophisticated strategies (TF2T, contrite TFT: in Langlois notes, defecting on guilty opponent maintains non guilty status - Markov Perfect, Subgame Perfect).
- Curiosity: TFT as extortionate? Any Subgame Perfect EQ has a countervailing and coercive part.
- Continuous case, replicator dynamics.
- Explain assumptions and conclusions in proofs rather than quote proofs.
- Example by Langlois.

# Chapter 1

## Abstract

One of the most famous models in game theory, the Prisoner's Dilemma, is an example of a game in which the Nash equilibrium solution is clearly not beneficial for either player. Regardless of the opponent's move, defection is always the dominant strategy in a single play of the game, and the resulting mutual defection costs both players the higher payoff they would have received from the unstable mutual cooperation play. If, however, the players have a nonzero chance of repeat interaction, turning the game into the Iterated Prisoner's Dilemma, defection may no longer be dominant, and the game allows for a rich field of strategies that utilize any available information about the opponent and vary their behavior between defection and cooperation (which can be viewed as punishment or reward) in the attempt to maximize the player's payoff or manipulate the opponent's payoff. Iterated Prisoner's Dilemma has been applied in ecology, sociology, evolutionary biology, international relations and behavior of firms in the market. This paper will review the history of Iterated Prisoner's Dilemma and its applications and will focus on the recent results by Press and Dyson (2012) showing the existence of a class of Zero Determinant strategies that allow a player to unilaterally determine her opponent's expected payoff, and to use extortion against some types of opponents. It will also examine the relationship between Zero Determinant strategies and the class of countervailing strategies proposed by J.P. Langlois and review the findings of Adami and Hintze, who show that the Zero Determinant strategies are not evolutionarily stable.

# Chapter 2

## Introduction

The question that inspired the body of research I focus on in this paper is simple to state: in a world of self-interested individuals (actors?), when does it make sense to cooperate? Real life examples abound. (*Include examples: Price fixing. Athletes and steroid use. Governments and trade. Cold war.*)

The question extends beyond human behavior as we observe patterns of cooperation everywhere in nature from reciprocal food exchange in vampire bats (expand), to countless examples interspecies cooperation or symbiosis (*Include examples in Axelrod, p90*).

These examples share a common structure. Each involves two participants making a choice to cooperate (fix prices, say no to steroids, remove trade restrictions, reduce nuclear weapons stocks, share blood with the less fortunate bat) or defect (undercut the competitor, use steroids, establish trade restrictions, increase weapons stockpiles, hoard the blood). Mutual cooperation yields a higher advantage to both parties than mutual defection, but if one individual cooperates, while the other defects, the defector gains a large advantage, while the cooperator is left a sucker.

This scenario was originally formalized in 1950 by Merrill Flood and Melvin Dresher at RAND corporation. Soon after Albert W. Tucker offered this illustrative parable that gave the game its name, the Prisoner's Dilemma. In Tucker's tale two prisoners are arrested on a minor charge. The prosecution suspects a more serious crime but does not have evidence to convict unless the prisoners testify against each other. Without the testimony (that is if the prisoners *cooperate* with each other), both would be convicted of the minor charge and receive a small sentence, say one month jail time. However if one prisoner testifies (or *defects* against the other), and the other prisoner keeps mum (cooperates), then the defecting prisoner will have the minor charge dropped, and the cooperating prisoner will receive a large sentence, say five years. Finally if both prisoners rat each other out (or mutually defect), they will both receive large, but lesser sentences, say two years each.

	Mum	Rat
Mum	(-0.1, -0.1)	(-5, 0)
Rat	(0, -5)	(-2, -2)

In most literature, however, the problem is presented with positive payoffs, usually with the following values.

	Cooperate	Defect
Cooperate	(3, 3)	(0, 5)
Defect	(5, 0)	(1, 1)

At first look the proposition may seem pessimistic. When the game is played once by players who have no possibility of future interaction, the only rational move is to defect even though mutual cooperation would lead to a higher payoff for both players. To see this we assume our opponent's move is fixed and analyze our choices. Given that our opponent cooperated, we may choose to cooperate and receive a payoff of 3, or defect for a payoff of 5, thus we choose to defect. If our opponent defected, then our choice of cooperation would yield 0, and defection would yield 1, thus regardless of the opponent's choice, it pays for us to defect. In other words if we choose the strategy of defection, we cannot improve our payoff by a unilateral change of strategy. This solution concept, one of the most famous and important in game theory, is called Nash equilibrium (NE) in honor of mathematician John Forbes Nash, who showed that at least one mixed strategy Nash equilibrium exists in any game with a finite set of actions [Wikipedia: [http://en.wikipedia.org/wiki/Nash\\_equilibrium#History](http://en.wikipedia.org/wiki/Nash_equilibrium#History)]

**Definition 2.1.** Nash equilibrium.

The game becomes more interesting—and less gloomy—with the possibility of repeat interaction, for it gives players a chance to reward cooperation and exact vengeance against defectors. When choosing a strategy, one must no longer just maximize the current payoff, but also consider the shadow of the future. Whereas in the singleton game the pure strategy space had only two points, the repeated game allows for great complexity and leads to some unexpected results that give insights about many problems in sociology, evolutionary biology, and political science.

The aim of this paper is to survey the history of the Iterated Prisoner's Dilemma focusing on four distinct approaches to its solution and their insights and consequences. In part one it will define and discuss the notion of evolutionary stability (ESS), due to John Maynard Smith, that paved the way for game theory applications to evolutionary biology and ecology. Then I will discuss the famous IPD tournaments hosted by Robert Axelrod and his conclusions about qualities shared by successful entrants. Part three will present an approach to modeling evolutionary systems using differential equations known as replicator dynamics. Finally part four will describe a class of Zero Determinant (ZD) strategies discovered by Press and Dyson in 2012 and discuss ZD strategies' evolutionary stability and role in mixed strategy populations.

# Chapter 3

## Preliminaries

Given a game, we may want to find the ‘best’ strategy, but first we need to develop some preliminaries in order to understand how to assess a strategy’s ‘goodness’. This section will define games in Normal and Extensive forms as well as pure and mixed strategies. It will also introduce rationality assumptions of different forms of Nash Equilibrium and briefly discuss the caveats of rationality assumptions.

**Definition 3.1.** A **normal form** game is a tuple  $G = (N, A, F)$ , where  $N = \{1, 2, \dots, m\}$  is the set of players,  $S = \{S_1, S_2, \dots, S_m\}$  is the set of actions, or pure strategies, one for each player, and  $F = \{F_1, F_2, \dots, F_m\}$  is a set of payoff functions,  $F_i: S_1 \times S_2 \times \dots \times S_m \rightarrow \mathbf{R}$ .

The description of the Prisoner’s Dilemma in the introduction is an example of a Normal form game. It specifies the set of players  $N = \{1, 2\}$ , the actions available to each  $S_1, S_2 \in \{\text{cooperate (C), defect (D)}\}$ , and the payoff function on the space of outcomes  $F: S_1 \times S_2 \rightarrow \mathbf{R} \times \mathbf{R}$  with  $F(C, C) = (R, R)$ ,  $F(C, D) = (S, T)$ ,  $F(D, C) = (T, S)$ , and  $F(D, D) = (P, P)$ , where the letters R, S, P, and T respectively stand for Reward (for mutual cooperation), Sucker (for cooperating while the opponent defects), P (punishment for mutual defection), and T (being a Traitor while the opponent cooperates). The values common in literature are  $T = 5$ ,  $R = 3$ ,  $P = 1$ , and  $S = 0$ , however any payoffs satisfying  $T > R > P > S$  maintain the Prisoner’s dilemma structure. A two-player game in normal form can be represented as a matrix of ordered pairs of payoffs, so it is sometimes called a bimatrix game. This form is useful for identifying Nash equilibria and eliminating strictly dominated strategies (Prisoner’s Dilemma does not have any strictly dominated strategies). Since the PD is symmetric, however, we can drop the opponents payoff from its representation and express it as a 2x2 matrix of row player’s payoffs.

	Cooperate	Defect
Cooperate	3	0
Defect	5	1

It is important to note that all of the game’s information is provided by its matrix, so players do not have a way to communicate their intentions prior to making the move, have no information about behavioral idiosyncrasies of the opponent, and derive no positive or negative utility from either spite (hurting their opponent), guilt (for hurting their opponent) or altruism (helping their opponent at a cost to themselves).

The normal form is convenient for eliminating strictly dominated strategies (which do not exist in the Prisoner's Dilemma), and finding Nash equilibria, however when we play the game repeatedly, the tree representation can better help us analyze the how the game unfolds in time.

**Definition 3.2.** (From Osborne and Rubinstein, p.90) An **extensive game with perfect information** has the following components.

1. A set  $N$  of players.
2. A set  $H$  of sequences (finite or infinite) that satisfies the following three properties.
  - (a) The empty sequence  $\emptyset$  is a member of  $H$
  - (b) If  $(a^k)_{k=1,2,\dots,K} \in H$  and  $L < K$  then  $(a^k)_{k=1,2,\dots,L} \in H$ .
  - (c) If an infinite sequence  $(a^k)_{k=1}^\infty$  satisfies  $(a^k)_{k=1,\dots,L} \in H$  for every positive integer  $L$ , then  $(a^k)_{k=1}^\infty \in H$ .

Each member of  $H$  is a **history**; each component of a history is an **action** taken by a player. A history  $(a^k)_{k=1,\dots,K} \in H$  is **terminal** if it is infinite or if there is no  $a^{K+1}$  such that  $(a^k)_{k=1,\dots,K+1} \in H$ . The set of terminal histories is denoted  $Z$ .

3. A function  $P$  that assigns to each nonterminal history (each member of  $H \setminus Z$ ) a member of  $N$ . ( $P$  is the **player function**,  $P(h)$  being the player who takes an action after the history  $h$ .)
4. For each player  $i \in N$  a preference relation  $\succeq_i$  on  $Z$ .

## Backward induction

We can assign every node a payoff value as defined by

Faced with a game, the player attempts to maximize her payoff. Since the outcome depends on all players' actions, the player must make certain rationality assumptions about her opponents. The key assumption is that the opponent is only interested in maximizing her own payoff.

Proposition 1, (p 15) in Axelrod is a simple argument that with high enough discount parameter, there is no single best strategy independent of opponent's strategy. ***This uses discounting, so needs to be either introduced in Axelrod section or after discounting is introduced in preliminaries.***

**Definition 3.3.** Game. Normal form game. Extensive form game.

Note that PD is not zero-sum, so the goal is not to beat the opponent but to do well for yourself, which is not the same thing.

**Definition 3.4.** Pure and mixed strategy.

**Definition 3.5.** Nash equilibrium.

Empty threats as motivation for subgame perfection.



**Definition 3.6.** Subgame Perfect Nash Equilibrium.

**Definition 3.7.** Generic game and equivalence. (This is used in Adami and Hintze, perhaps makes more sense to define it right before using in the ZD section of the paper).

# Chapter 4

## Axelrod's Tournaments and TFT

*this section's text is not the final text, just a roadmap*

Axelrod held two rounds of tournaments. The first with 6? submissions, the second with 62. TFT won both tournaments. Axelrod analyzes the success of TFT. In particular he shows two other strategies that could have beat TFT in the first tournament. These 'better' strategies were among entrants in the second tournament, but did not do very well as they were thwarted by other new entrants. It's interesting to see that many strategies that did not end up doing well for themselves were important in creating a fitness landscape for other strategies, that is they had large effect on scores of other strategies, but did not have a strong effect on TFT - an argument for TFT's robustness.

Then Axelrod motivates the evolutionary approach with average payoffs of each tournament (generation) determining reproductive fitness. With the starting set of strategies of the second tournament, TFT prevails, that success is robust to changing the initial distribution of strategies. This provides the segue to ESS section that follows.

Discussion of *collective stability* will probably be omitted, as it's almost the same as ESS. A strategy is collectively stable if it is in Nash Equilibrium with itself. ESS are collectively stable. TFT is collectively stable, while the only ESS is AllD. Strategies that invade AllD the best are called *maximally discriminating* (p. 66).

Can a population of TFT be invaded? Proposition 7, p 67 shows that if a nice strategy cannot be invaded by a single individual, then it cannot be invaded by a cluster.

This section will motivate the discounted IPD model by showing that discounting is equivalent to a fixed probability of repeat encounter. Introduce TFT and describe its four characteristics that Axelrod concluded were key to its success (*these are taken from Wikipedia, will find exact words in Axelrod later*):

- Be nice: cooperate, never be the first to defect.
- Be provokable: return defection for defection, cooperation for cooperation.
- Don't be envious:: be fair with your partner.
- Don't be too clever: or, don't try to be tricky.

Discuss TFT and ESS (TFT is ESS if and only if the discount factor is sufficiently great). TFT does not earn more than opponent. TFT's success is in its ability to do well both

against ‘nice’ and against ‘bad’ strategies. Discuss how TFT variants did not perform as well as TFT (JOSS - TFT with a probabilistic opportunistic defection vs. TFT example, p 39). Small group of TFT can invade ALLD populations. Population can be split into different subpopulations that are TFT to members of the same subpopulation, but ALLD to the other subpopulation.

Importance of being nice. JOSS v TFT showed us that opportunistic defection hurt the strategy’s score in the tournament. Axelrod shows that TFTT (TIT FOR TWO TATS) would have actually won the tournament if entered. He notes that most submissions were trying to gain advantage from opportunistic defection, but the bigger advantage could be won by instead being more forgiving/nicer.

### 4.0.3 Examples

- ‘Live and let live’ cooperation in the trench warfare during WWI. Enabled by the static nature of the war when soldiers faced the same soldiers on the other side for long periods of time. Soldiers and officers ignored the orders of the high command and only fired warning/demonstration shots.
- There is a list of references to papers about arms race, oligopolistic competition and vote trading on (p. 28), and non-empirical papers on meaning of rationality (Luce and Raiffa), choices which affect other people (Schelling), Cooperation without enforcement (Taylor) (p. 29)

### 4.0.4 Other points and themes in Axelrod

***The following quote perhaps better used or paraphrased in the introduction to motivate interest in IPD.*** “The model of the iterated Prisoner’s Dilemma is much less restricted than it may at first appear. Not only can it apply to interactions between two bacteria or interactions between two primates, but it can also apply to the interactions between a colony of bacteria and, say, a primate serving as a host. There is no assumption that payoffs of the two sides are comparable. Provided that the payoffs to each side satisfy the inequalities that define the Prisoner’s Dilemma, as given in chapter 1, the results of the analysis will be applicable” (p 95)

***The following quote may be used in ESS section or Axelrod section*** “The chronological story that emerges from this analysis is the following. ALL D is the primeval state and is evolutionarily stable. But cooperation based on reciprocity can gain a foothold through two different mechanisms. First, there can be kinship between mutant strategies, giving the genes of the mutants some stake in each other’s success, thereby altering the payoff of the interaction when viewed from the perspective of the gene rather than the individual. A second mechanism to overcome ego tap defection is for the mutant strategies to arrive in a cluster so that they provide a nontrivial proportion of the interactions each has, even if they are so few as to provide a negligible proportion of the interactions which the ALL D individuals have. Then the tournament approach described in chapter 2 demonstrates that once a variety of strategies is present, TIT FOR TAT is an extremely robust one. It does well in a wide range of circumstances and gradually displaces all other strategies in an ecological

simulation that contains a great variety of more or less sophisticated decision rules... *Thus cooperation based on reciprocity can get started in a predominantly noncooperative world, can thrive in a variegated environment, and can defend itself once fully established* (emphasis added).” (p. 99)

***Mechanisms for mutualism when players can’t recognize each other.*** Continuous contact (hermit crab & sea-anemone, tree and its mycorrhizal fungi); constant meeting place (aquatic cleaner mutualizes in coastal and reef situations (Trivers 1971). (p 101). “Territory can serve this purpose.... Consistent with the theory, such male territorial birds show much more aggressive reactions when the song of an unfamiliar male rather than a neighbor is reproduced nearby” (p. 102)

***Perhaps a curious side note.*** Prosopagnosia and the location of lesions that cause it: “This localization of cause, and specificity of effect, indicates that the recognition of individual faces has been an important enough task for a significant portion of the brain’s resources to be devoted to it” (p. 102)

***Another side note.*** There have been papers that applied IPD to model the increase in pathogen activity when the host is weakened (possibly by another pathogen), which could increase the chances that the first pathogen can transmit. (p. 103-104).

***Another side note.*** Cooperation is not necessarily a socially desirable outcome. Examples: corruption is an act of cooperation between politicians and moneyed interests, price fixing is an act of cooperation between companies. (p. 18)

***Another side note.*** “Examples of what is left out by this formal abstraction include the third parties, the problems of implementing a choice, and the uncertainty about what the other player actually did on the preceding move.” (p. 19)

#### 4.0.5 Beating TFT: Pavlov

The success of IPD and TFT is at least partially to their simplicity – simpler models are more likely to give insight to a large number of problems than more complicated ones (analogy to bias-variance tradeoff, more specific models are prone to higher variance/overfitting *probably won’t use this analogy, but will keep it here for now*). Can we do better than TFT while preserving model parsimony? In 1993 Nowak and Sigmund published a paper in Nature describing Pavlov, the strategy of WIN-STAY, LOSE-SHIFT.

##### Some points from Pavlov paper

- Keeps the advantage of model parsimony.
- “Pavlov seems to be more robust than tit-for-tat, suggesting that cooperative behavior in natural situations may often be based on win-stay, lose-shift”.
- “The conspicuous success of the tit-for-tat strategy relies in part on the clinical neatness of a deterministic cyber-world. In natural populations, errors occur. TFT suffers from stochastic perturbations in two ways: (1) a TFT population can be ‘softened up’ by random drift introducing unconditional cooperators, which allow exploiters to grow (TFT is not an evolutionarily stable strategy *we already showed that TFT is ESS with sufficiently high discount factor, need to see what they mean*

*here*); and (2) occasional mistakes between two TFT players cause long runs of mutual backbiting. (Such mistakes abound in real life: even humans are apt to vent frustrations upon innocent bystanders.)”

- “Within the restricted world of strategies reacting only to the co-players previous move, TFT has a very important, but transitory role: in small clusters, it can invade populations of defectors, but then bows out to a related strategy, ‘generous tit for tat’ (GTFT), which cooperates after a co-player’s C, but also with certain probability after a D... but as soon as one admits strategies which take into account the moves of both players in the previous round, evolution becomes much less transparent. We first conjectured that GTFT (or variants thereof) would win the day, but are forced to admit, after extensive simulations, that the strategy Pavlov did much better in the long run.”
- the paper uses infinitely iterated PD ( $w = 1$ )
- Pavlov can be invaded by AllD if  $2R < T + P$ , but a prudent variant of it  $(1, 0, 0, x)$  with  $x < (R - P)/(T - R)$ , is stable against AllD. For  $2R = T + P$  (as is the case with Axelrod’s payoff values), Pavlov can be invaded by AllD, but the stochastic, Pavlov-like strategy  $(0.999, 0.001, 0.001, 0.995)$  cannot.
- “How does this relate to real biology? It might as well be that many cases of cooperation based on reciprocal altruism are due to Pavlov rather than to TFT. In more natural set-ups, retaliation has usually been interpreted as evince for TFT. But these experiments seem to be consistent with a Pavlov-like strategy as well. It would speak for Pavlov if animals have a tendency to exploit non-retaliators, or if they are apt to resume cooperation after bilateral defection.”

# Chapter 5

## ESS

This section will contain motivation of ESS in terms of stability and resistance to invasion, formal definition using  $\epsilon$ -neighborhoods. Based on Maynard Smith's article and book *The Logic of Animal Conflict* and *Evolution and the Theory of Games*.

### 5.1 This quote can be used or paraphrased as a segue into ESS

“Paradoxically, it has turned out that game theory is more readily applied to biology than to the field of economic behavior for which it was originally designed. There are two reasons for this. First, the theory requires that the values of different outcomes (for example, financial rewards, the risks of death and the pleasures of a clear conscience) be measured on a single scale. In human applications, this measure is provided by ‘utility’ – a somewhat artificial and uncomfortable concept: in biology, Darwinian fitness provides a natural and genuinely one-dimensional scale. Secondly, and more importantly, in seeking the solution of a game, the concept of human rationality is replaced by that of evolutionary stability. The advantage here is that there are good theoretical reasons to expect populations to evolve to stable states, whereas there are grounds for doubting whether human beings always behave rationally” (Maynard Smith, Preface)

### 5.2 Body of ESS section

“A ‘strategy’ is a behavioral phenotype; i.e. it is a specification of what an individual will do in any situation in which it may find itself... the concept is couched in terms of a ‘strategy’ because it arose in the context of animal behavior. The idea, however, can be applied equally well to any kind of phenotypic variation, and the word strategy could be replaced by the word phenotype; for example, a strategy could be the growth form of a plant, or the age at first reproduction, or the relative numbers of sons and daughters produced by a parent” (p. 10)

## 5.3 TFT is ESS

Proof that TFT is an ESS using the idea that if a strategy is a best response to TFT, then it must make the same decision every time it encounters the initial state, so only AllC, AllD and DCDC... are possible best responses, and for large enough discount factors they do worse than TFT. (Maynard Smith p. 203)

## 5.4 TFT is not ESS after all?

In article *No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game* (Nature, 1987) Boyd and Lorberbaum show that TFT is not ESS. I haven't yet worked through their argument.

History of game theoretic modeling of evolution: Lewontin (1961), Fisher (1930), Hamilton (1967) (p.2)

Stability of AllD.

## 5.5 Quotes and other points from Maynard Smith

"An obvious weakness of the game-theoretic approach to evolution is that it places great emphasis on equilibrium states, whereas evolution is a process of continuous, or at least periodic change. The same criticism can be leveled at the emphasis on equilibria in population genetics... there are, however, two situations in which game theory models force us to think about change as well as constancy." these situations are not reflected in IPD (when either the game has more than one or no ESS, IPD's only ESS is AllD). (p. 8)

On truth or testability of game theoretic models: "I think it would be a mistake, however, to stick too rigidly to the criterion of falsifiability when judging theories in population biology. For example, Volterra's equations... their merit is to show that even the simplest possible model of such an interaction leads to sustained oscillation – a conclusion it would have been hard to reach by purely verbal reasoning."

**Segue to the next section:** the ESS gives us both an intuitive framework and a formal topologic understanding of fitness of strategies. The proliferation of cheap computing power in the 1970s gave us another tool to study strategies – agent-based models.

### Another example of non-PD ESS

Egg recognition vs. altruistic adoption in Guillemots as an example of ESS in nature.  
*Dawkins, p.102*

# Chapter 6

## Preliminaries

Some preliminary definitions and general results are in order before we immerse ourselves in

### 6.1 Games and Solution concepts

This may be split up in subsection or just presented as a series of definitions with some motivation.

#### 6.1.1 Games and Iterated Games

Normal and extensive form. Prisoner's dilemma. Repeated games and discounting. Continuous PD (if I discuss countervailing strategies).

**Definition 6.1.** Game.

**Definition 6.2.** Normal and extensive form games.

**Example 6.1.** Prisoner's dilemma, coordination game (or Stag Hunt; appears in Zero Determinant with Tags analysis in Adami and Hintze), Hawk-Dove (may exclude).

**Definition 6.3.** Generic game and equivalence. (Gintis, p.285)

“Suppose a normal form game is generic in the sense that no two payoffs for the same player are equal. Suppose  $A = (a_{ij})$  and  $B = (b_{ij})$  are the payoff matrices for Alice and Bob, so the payoff to Alice's strategy  $s_i$  against Bob's strategy  $t_j$  is  $a_{ij}$  for Alice and  $b_{ij}$  for Bob. We say that two generic  $2 \times 2$  games with payoff matrices  $(A, B)$  and  $(C, D)$  are equivalent if”, for all  $i, j = 1, 2$

$$a_{ij} > a_{kl} \equiv c_{ij} > c_{kl}$$

and

$$b_{ij} > b_{kl} \equiv d_{ij} > d_{kl}$$

**Proposition 6.1.** “Every generic  $2 \times 2$  game is equivalent to either the prisoner's dilemma, the battle of the sexes, or the hawk-dove.” (Gintis, p.286)



Karl Sigmund in Calculus of Selfishness defines average payoff in addition to the geometric discounting with  $\mathbb{P}(\text{another iteration}) = w < 1$  which is useful for the limiting case  $w = 1$ :

$$\lim_{n \rightarrow \infty} \frac{A(0) + \cdots + A(n)}{n + 1}$$

where  $A(n)$  is the payoff in round  $n$ .

### 6.1.2 Nash equilibrium

**Definition 6.4.** Nash equilibrium, strict Nash equilibrium.

**Theorem 6.2.** *Every symmetric game admits a symmetric Nash equilibrium. (Sigmund, Section 2.5)(May leave out)*

### 6.1.3 Subgame perfection

TFT is not subgame perfect since if a player decides to defect on a certain move and then return to TFT against a TFT opponent, then they will continue DCDC... cycle with a lower payoff than CCC... (Gintis, p. 209). Contrite TFT is subgame perfect.

### 6.1.4 Markov perfection

If I discuss countervailing strategies.

### 6.1.5 ESS

May include criticisms of John Maynard Smith's ESS definition and discuss more modern alterations. Alternatively, the criticisms and discussion may be included in the discussion of papers.

**Definition 6.5.**  $x \in \Delta$  is an evolutionary stable strategy (ESS) if for every strategy  $y \neq x$  there exists some  $\bar{\epsilon}_y \in (0, 1)$  such that inequality

$$u[x, \epsilon y + (1 - \epsilon)x] > u[y, \epsilon y + (1 - \epsilon)x]$$

holds for all  $\epsilon \in (0, \bar{\epsilon}_y)$ .

**Proposition 6.3.**

$$\Delta^{ESS} = \{x \in \Delta^{NE} : u(y, y) < u(x, y) \ \forall y \in \beta^*(x), y \neq x\}$$

“Also if  $(x, x) \in \Theta$  is a strict Nash equilibrium, then  $x$  is evolutionarily stable by default—then there are no alternative best replies. This observation has immediate implications concerning the connection between evolutionary stability and social efficiency: Evolutionary stability does not in general imply that average population fitness  $u(x, x)$  is maximized.” (Weibull, P38).

An example of potential social inefficiency of ESS is the one-shot prisoner's dilemma, whose only NE, and thus only ESS is to always defect.

There are games with no ESS, for example Rock Paper Scissors (Weibull, P.40)

May include results about ESS and trembling hand perfection (Weibull, P.42). Evolutionary stability requires behavior that is not only "rational" and "coordinated" in the sense of Nash equilibrium but also "cautious."

## Invasion barriers

In the "setting of finite games, evolutionary stability implies that  $\bar{\epsilon}_y$  can be taken to be the same for all mutants; that is, an evolutionary stable strategy  $x$  has a uniform invasion barrier." (Weibull, P.43)

**Proposition 6.4.**  $x \in \Delta^{ESS}$  if and only if  $x$  has a uniform invasion barrier. (Weibull, P.43).

### 6.1.6 Some strategies

TFT, Generous TFT, Pavlov (Win-Stay, Lose-Shift).

## 6.2 Replicator dynamics

This section will be based either on *Game Theory Evolving* (Gintis, H.), *Evolutionary Game Theory* (Weibull, J.W.), or *The Calculus of Selfishness* (Sigmund, K).

**Definition 6.6.** A replicator dynamic.

**Definition 6.7.** Fixed point.

Will possibly have some discussion of imitation dynamics, but for now it looks like imitation is not used in any of the papers I am discussing, so I will most likely leave out imitation.

**Theorem 6.5.** *In symmetric  $2 \times 2$  games a population state is asymptotically stable in the replicator dynamics if and only if the corresponding mixed strategy is evolutionarily stable. (Weibull, p.75).*

**Theorem 6.6.** *Strongly dominated strategies do not survive in a replicator dynamic. (Theorem 12.3, Gintis, p.280). Stated without proof.*

**Theorem 6.7.** *Weakly dominated strategy cannot achieve unitary probability as  $t \rightarrow \infty$  in a replicator dynamic. (Theorem 12.4, Gintis, p.281). Stated without proof.*

## **Replicator dynamics and ESS**

There is an example of a Nash equilibrium (Gintis, section 10.13) that “cannot be invaded by any pure strategy mutant but can be invaded by an appropriate mixed-strategy mutant. We can show that this Nash equilibrium is unstable under the replicator dynamic. This is why we insisted that the ESS concept be defined in terms of mixed- rather than pure-strategy mutants; an ESS is an asymptotically stable equilibrium only if the concept is so defined.” (Gintis, p. 286)

## **6.3 Agent based modeling**

I’m not yet sure I want to discuss this approach in the preliminaries. It seems pretty self-explanatory, so I may just refer to it when citing and discussing results.

## **6.4 Learning**

A short description of learning. Press and Dyson talk about performance of extortionate ZD strategies against an opponent who learns by climbing up his payoff gradient.

# Chapter 7

## Results

I haven't yet decided if each of the papers I focus on will occupy its own section or if their results will be woven in a more continuous narrative. I may revisit Maynard Smith's ESS here.

### 7.1 Axelrod, tournaments, TFT

Axelrod's characterization of successful competitors (nice, retaliating, forgiving).

### 7.2 Press and Dyson: zero determinant strategies

This will be a summary of the paper's arguments along with the proofs of the theorems. The three results are:

**Theorem 7.1.** *There exists strategies that can unilaterally set the opponent's score or demand and get an extortionate share. Press and Dyson note that the ability to unilaterally set the opponent's score allows the ZD player to simulate an arbitrary fitness landscape for the evolutionary opponent. They also discuss what happens when extortionate ZD player plays against an evolutionary opponent.*

**Theorem 7.2.** *Shortest memory player sets the rules of the game.*

*(Note from Adami & Hintze: this is true in a head-to-head competition, however in evolutionary setting longer memory may offer an advantage as the strategy would be able to recognize an opponent playing the same strategy, which would enable the long-memory player to conditionally cooperate, a conditionally cooperating ZD can be ESS).*

**Theorem 7.3.** *ZD strategies succeed without Markov equilibrium. Press and Dyson used the game's stationary distribution to derive their ZD strategies. This result shows that the opponent cannot "keep the game out of Markov equilibrium" or play "inside the Markov equilibration time scale."*

There is another proof of existence of ZD strategies in (Hilbe, Nowak, Sigmund, 2013), but I will most likely use the Press and Dyson proof.

### 7.2.1 Stewart and Plotkin comments on Press and Dyson

Stewart and Plotkin ran an Axelrod-style tournament using the usual set of strategies and two additional ZD strategies Extort-2 ( $S_X - P = 2(S_Y - P)$ ) and Zero Determinant Generous Tit For Tat, ZDGTFT-2 ( $S_X - R = 2(S_Y - R)$ ) and found that Extort-2 had the second number of head-to-head wins (the winningest strategy, of course, was the Pyrrhic AllD), while ZDGTFT-2 achieved the highest score.

## 7.3 Evolutionary stability of ZD

This section will be based on two papers. Their abstracts are.

### 7.3.1 Adami and Hintze, 2013

“Here we show that ZD strategies are at most weakly dominant, are not evolutionarily stable, and will instead evolve into less coercive strategies. We show that ZD strategies with an informational advantage over other players that allows them to recognize each other can be evolutionarily stable (and able to exploit other players). However, such an advantage is bound to be short-lived as opposing strategies evolve to counteract the recognition”.

Adami and Hintze define a new game in which players chose a role in  $\{ZD, O\}$  (I think this is for ZD that are not extortionate), and the payoffs are the long-term average payoffs when playing an IPD with chosen roles. They then analyze this new game using replicator equations. In particular they look at two-strategy populations  $\{AllD, ZD\}$  and  $\{Pavlov, ZD\}$ . ZD loses to AllD in evolutionary contest. Pavlov,  $q_{PAV} = (1, 0, 0, 1)$ , cooperates with itself, is ESS, and loses to ZD in every direct competition, however Pavlov drives ZD to extinction (using replicator equations).

Then they acknowledge that the reformulation of the game may not perfectly reflect the dynamics that happen when the agents in the population play random one-shot games against each other. To address that, they created agent-based simulations and found that

### 7.3.2 Hilbe, Nowak, Sigmund, 2013

“Here, we analyze the evolutionary performance of this new class of strategies. We show that in reasonably large populations, they can act as catalysts for the evolution of cooperation, similar to TFT, but that they are not the stable outcome of natural selection. In very small populations, however, extortioners hold their ground. Extortion strategies do particularly well in coevolutionary arms races between two distinct populations. Significantly, they benefit the population that evolves at the slower rate, an example of the so-called “Red King” effect. This may affect the evolution of interactions between host species and their endosymbionts”.

## 7.4 Countervailing

# Chapter 8

## Other results and discussions

Some of these may make more sense to include in the above discussions than give them their own section. Some may just not be a good fit for this paper.

### 8.1 An empirical approach to PD

From Lave, L.B. 1962. There is evidence that people do not play the Nash Equilibrium even in iterated prisoner's dilemma with a fixed number of plays. In this setting . These results could be added to the solution concepts chapter.

### 8.2 Conclusion

From *The Selfish Gene*, 3rd Ed, p.75:

This theoretical conclusion is not far from what actually happens in most wild animals. We have in a sense explained the 'gloved fist' aspect of animal aggression. Of core the details depend on the exact numbers of 'points' awarded for winning, being injured, wasting time, and so on. In elephant seals the prize for winning meat be near-monopoly rights over a large harem of females. The pay-off for winning must therefore be rated very high. Small wonder that fights are vicious and the probability of serious injury is also high. The cost of wasting time should presumably be regarded as small in comparison with the cost of being injured and the benefit of winning. For a small bird in a cold climate, on the other hand, the cost of wasting grime may be paramount. A great tit when feeding nestlings needs to catch an average of one prey per thirty seconds. Every second of daylight is precious. Even the comparatively short time wasted in a hawk/hawk fight should perhaps be regarded as more serious than the risk of injury to such a bird. Unfortunately we know too little at present to assign realistic numbers to the costs and benefits of various outcomes in nature\*...

## 8.3 Notation

Notation from various sources. Once the content of the paper is clear, I will make sure notation is consistent.

$\Delta$  strategy space (Weibull)

$\Theta$  strategy profile (Weibull)