Essays on nonlinear evolutionary game dynamics
Ochea, M.I.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Evolutionary game theory has been viewed as an evolutionary “repair” of rational actor game theory in the hope that a population of boundedly rational players may attain convergence to classic “rational” solutions, such as the Nash Equilibrium, via some learning or evolutionary process.

In this thesis the model of boundedly rational players is a perturbed version of the best-reply choice, the so-called Logit rule.

With the strategic context varying from models of cyclical competition (Rock-Paper-Scissors), through industrial organization (Cournot) and to collective-action choice (iterated Prisoner’s Dilemma), we show that the logit evolutionary selection among boundedly rational strategies does not necessarily guarantee convergence to equilibrium and richer dynamical behavior - e.g. cycles, chaos - may be the rule rather than the exception.

Marius-Ionut Ochea (1977) holds a M.A. in Economics from Central European University, Budapest, Hungary (2004) and a M.Phil. in Economics from Tinbergen Institute, Amsterdam, the Netherlands (2006). In 2006 he joined CeNDEF (Center for Nonlinear Dynamics in Economics and Finance) at the University of Amsterdam for his Ph.D. study. As of September 2009 he is working as a post-doc researcher at CentER, Tilburg University. Main research interests include: evolutionary game dynamics, norms evolution and emergence of cooperative behavior.
Essays on Nonlinear Evolutionary Game Dynamics
ISBN 978 90 5170 688 8

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 468 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute.

A list of books which already appeared in the series can be found in the back.
Essays on Nonlinear Evolutionary Game Dynamics

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promotie ingestelde commissie, in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 19 januari 2010, te 14.00 uur

door

Marius Ionut Ochea

geboren te Vaslui, Roemenië
Promotiecommissie:

Promotor: Prof. Dr. Cars H. Hommes

Overige leden: Prof. Dr. Josef Hofbauer
               Prof. Dr. Aart de Zeeuw
               Dr. Roald Ramer
               Dr. Jan Tuinstra

Faculteit Economie en Bedrijfskunde
Acknowledgements

It may not come as great surprise to say that, some 5 years ago, when I was joining the Tinbergen Institute with the purpose of pursuing doctoral research, I had only limited insight into what the actual topic of my dissertation would eventually look like. I would have probably found some words in the current title of this thesis very strange and exotic. Beyond a broad interest in the scientific adventure, the concrete pathmarks were yet to be determined. TI M.Phil first year programme proved flexible enough to allow me roam around the institute academic diversity and make a first contact with CeNDEF researchers during lectures on nonlinear dynamics, bounded rationality or evolutionary game theory. From the outset, I want to express my entire gratitude to Cars, Jan and Roald for awakening my interest towards these research paths. Later on, I was glad to apply to a PhD position at CeNDEF and start working on a project that combines exactly these three lines of investigation. My dissertation emerged out of this process.

It is perhaps difficult to strike the right tone when looking back in time at people that marked my experience as a doctoral student. First, it was a great privilege to have Prof. Cars Hommes as my supervisor: not only for the role his advice played on structuring and disciplining my pretty chaotic way of doing research but also his constant support and encouragements throughout the thesis-writing process. His friendly attitude, dedication, extensive feedback and minutely reading of my earlier drafts were decisive for the timely and successful completion of the thesis.

I am also thankful to the members of my thesis committee. I had stimulating
conversations with Prof. Josef Hofbauer on earlier versions of Chapter 2 and some extensions in that chapter originate from his suggestions. Jan Tuinstra commented at length on Chapter 3 whereas Chapter 4 was initiated out of our discussions. Roald Ramer introduced me to the field of evolutionary game dynamics and we had innumerable exchanges of ideas on various applications of the evolutionary theory ranging from language to preferences evolution. Aart de Zeeuw gave me the opportunity to continue work on repeated Prisoner’s Dilemma done in Chapter 5 and engage in post-doctoral work on the emergence and evolution of institutions sustaining cooperation at the international level.

At various stages of my research I have also benefited a lot from interaction and very kind advise from a scientifically heterogenous group of researchers as CeNDEF is: Dave, Cees, Florian, Pim, Maurice, Misha, Stelios. For the lively and friendly environment I enjoyed in the group, my gratitude goes also to former or current CeNDEF doctoral students: Pietro, Valentyn, Peter, Saeed, Tatiana, Domenico, Paolo, Te. Officemates are a sort of a peculiar species in research and I would like to thank Pim, not only for our almost daily debates on almost any subject one could think of, but also for his virtually unlimited availability for exploring the local (experimental) music scene at, among others, STEIM or DNK Amsterdam. Outside CeNDEF, I would also want to mention former colleagues and friends from Tinbergen Institute: Antonio, Sumendhu, Alex, Aufa, Robert, Ana, Sumedha, Sebi, Razvan, Vali, Marcel, Mario, Jonneke, Tse-Chun. They all turned my early transition and stay in Amsterdam into smooth and pleasant experiences.

I could not complete this short review without thanking my parents for their patience and understanding towards my always too rare and too short trips home since my departure abroad. This thesis is heartly dedicated to them.

Amsterdam, November 2009
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contents</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiii</td>
</tr>
</tbody>
</table>

## 1 Introduction

1.1 Literature on convergence of game dynamics                           4
1.2 Literature on complicated game dynamics                               5
1.3 Thesis Outline                                                        7
    1.3.1 Multiple Steady States, Limit Cycles and Chaotic Attractors in Logit Dynamics 8
    1.3.2 Heterogenous Learning Rules in Cournot Games                      10
    1.3.3 On the Stability of the Cournot Solution: An Evolutionary Approach
    1.3.4 Evolution in Iterated Prisoner’s Dilemma Games under Smoothed Best-Reply Dynamics 12

## 2 Multiple Steady States, Limit Cycles and Chaotic Attractors in Logit Dynamics

2.1 Introduction                                                          15
    2.1.1 Motivation                                                        15
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.2 ‘Replicative’ vs. ‘rationalistic’ dynamics</td>
<td>18</td>
</tr>
<tr>
<td>2.2 The Logit Dynamics</td>
<td>19</td>
</tr>
<tr>
<td>2.2.1 Evolutionary dynamics</td>
<td>19</td>
</tr>
<tr>
<td>2.2.2 Discrete choice models—the Logit choice rule</td>
<td>21</td>
</tr>
<tr>
<td>2.3 Hopf and degenerate Hopf bifurcations</td>
<td>23</td>
</tr>
<tr>
<td>2.4 Three strategy games</td>
<td>26</td>
</tr>
<tr>
<td>2.4.1 Rock-Scissors-Paper Games</td>
<td>27</td>
</tr>
<tr>
<td>2.4.2 Coordination Game</td>
<td>39</td>
</tr>
<tr>
<td>2.4.3 Rock-Scissors-Paper and wLogit Dynamics</td>
<td>52</td>
</tr>
<tr>
<td>2.4.4 Schuster et al.(1991) Game and wLogit Dynamics</td>
<td>55</td>
</tr>
<tr>
<td>2.6 Conclusions</td>
<td>58</td>
</tr>
<tr>
<td>2.7 Rock-Scissor-Paper game with Logit Dynamics: Computation of the first Lyapunov coefficient</td>
<td>59</td>
</tr>
<tr>
<td>3 Heterogenous Learning Rules in Cournot Games</td>
<td>61</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>61</td>
</tr>
<tr>
<td>3.2 Standard Static Cournot Analysis</td>
<td>64</td>
</tr>
<tr>
<td>3.3 Heterogenous Learning Rules</td>
<td>65</td>
</tr>
<tr>
<td>3.3.1 Adaptive Expectations</td>
<td>65</td>
</tr>
<tr>
<td>3.3.2 Fictitious Play</td>
<td>66</td>
</tr>
<tr>
<td>3.3.3 Weighted Fictitious Play</td>
<td>66</td>
</tr>
<tr>
<td>3.4 Evolutionary Cournot Games</td>
<td>67</td>
</tr>
<tr>
<td>3.4.1 Adaptive Expectations vs. Rational/Nash play</td>
<td>67</td>
</tr>
<tr>
<td>3.4.2 Adaptive vs. Exponentially Weighted Fictitious Play</td>
<td>76</td>
</tr>
<tr>
<td>3.4.3 Local stability analysis</td>
<td>78</td>
</tr>
<tr>
<td>3.4.4 Naive vs. Fictitious Play</td>
<td>84</td>
</tr>
<tr>
<td>3.5 Concluding Remarks</td>
<td>87</td>
</tr>
</tbody>
</table>
# On the Stability of the Cournot Solution: An Evolutionary Approach

4.1 Introduction .................................................. 89
4.2 The model ................................................... 91
4.3 Results ......................................................... 96
   4.3.1 Best-response dynamics limit, $\beta \to \infty$ .................. 97
   4.3.2 Costly Rational Expectations, $k > 0$ and finite $\beta$ .......... 98
4.4 Conclusions .................................................. 101

# Evolution in Iterated Prisoner’s Dilemma Games under Smoothed Best-Reply Dynamics

5.1 Introduction .................................................. 103
5.2 An Evolutionary Iterated PD game ............................. 105
5.3 2×2 Ecologies .................................................. 109
   5.3.1 AllD vs. TFT ........................................... 109
   5.3.2 TFT vs. AllC ........................................... 111
   5.3.3 AllD vs. AllC ........................................... 114
   5.3.4 AllD vs. GTFT .......................................... 115
   5.3.5 AllD-WSLS ............................................. 117
   5.3.6 GTFT vs. AllC .......................................... 120
   5.3.7 GTFT vs. WSLS ......................................... 121
   5.3.8 TFT vs. GTFT ........................................... 124
   5.3.9 TFT vs. WSLS ........................................... 127
   5.3.10 WSLS vs. AllC ......................................... 128
   5.3.11 Summary ................................................ 130
5.4 3×3 Ecologies .................................................. 132
   5.4.1 AllD-TFT-AllC .......................................... 132
   5.4.2 AllD-GTFT-WSLS ..................................... 135
   5.4.3 AllD-GTFT-AllC ..................................... 139
5.4.4 AllD-TFT-WSLS ................................................. 141
5.4.5 AllD-TFT-GTFT ............................................... 144
5.4.6 AllD-WSLS-AllC ............................................... 146
5.4.7 TFT-WSLS-AllC ............................................... 148
5.4.8 TFT-GTFT-WSLS ............................................... 150
5.4.9 Summary ..................................................... 152

5.5 4×4 ecologies .................................................. 153
5.5.1 No TFT ...................................................... 153
5.5.2 No AllD ...................................................... 156
5.5.3 No GTFT ...................................................... 159
5.5.4 No WSLS ..................................................... 161
5.5.5 No AllC ...................................................... 163
5.5.6 Summary ..................................................... 165

5.6 5×5 Ecology .................................................... 166
5.6.1 Numerical Bifurcation Curves .............................. 170
5.7 Conclusions .................................................... 172
5.8 Iterated PD Game-stationary distributions and average payoffs . . . . . . . 173

6 Summary ......................................................... 181

Bibliography ....................................................... 185

Samenvatting - summary in Dutch ................................ 195
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Coordination Game, Replicator Dynamics–Long-run average fitness</td>
<td>42</td>
</tr>
<tr>
<td>4.1</td>
<td>Linear $n$-player Cournot game with costly Rational expectations - In-</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>stability Thresholds</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>Summary dynamics 2x2 ecologies</td>
<td>130</td>
</tr>
<tr>
<td>5.2</td>
<td>Summary dynamics 3x3 ecologies</td>
<td>153</td>
</tr>
<tr>
<td>5.3</td>
<td>Summary dynamics 4x4 ecologies</td>
<td>165</td>
</tr>
<tr>
<td>5.4</td>
<td>Stationary distributions and average payoffs for IPD</td>
<td>178</td>
</tr>
</tbody>
</table>
### List of Figures

<table>
<thead>
<tr>
<th>Figure Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Rock-Scissors-Paper and Replicator Dynamics</td>
<td>31</td>
</tr>
<tr>
<td>2.2</td>
<td>RSP Game, Logit Dynamics-Hopf scenario, $\beta$</td>
<td>33</td>
</tr>
<tr>
<td>2.3</td>
<td>RSP Game, Logit Dynamics-Hopf scenario, $\delta$</td>
<td>34</td>
</tr>
<tr>
<td>2.4</td>
<td>Rock-Scissors-Paper and Logit Dynamics: Hopf curves</td>
<td>35</td>
</tr>
<tr>
<td>2.5</td>
<td>Generalized Rock-Scissors-Paper - curves of LP and H bifurcitations</td>
<td>38</td>
</tr>
<tr>
<td>2.6</td>
<td>Coordination Game and Replicator Dynamics - Invariant manifolds</td>
<td>40</td>
</tr>
<tr>
<td>2.7</td>
<td>Coordination Game and Replicator Dynamics - Basins of attraction</td>
<td>42</td>
</tr>
<tr>
<td>2.8</td>
<td>Coordination Game, $\varepsilon = 0.1$ and Logit Dynamics - Multiple steady states</td>
<td>44</td>
</tr>
<tr>
<td>2.9</td>
<td>Coordination Game, $\varepsilon = 0.1$ and Logit Dynamics - Moderate $\beta$</td>
<td>45</td>
</tr>
<tr>
<td>2.10</td>
<td>Coordination Game, Logit Dynamics - Equilibria curves</td>
<td>47</td>
</tr>
<tr>
<td>2.11</td>
<td>Coordination Game, Logit Dynamics - Limit point curves</td>
<td>48</td>
</tr>
<tr>
<td>2.12</td>
<td>Coordination Game ($\varepsilon = 0.1$) and Logit Dynamics–Basins of Attraction</td>
<td>49</td>
</tr>
<tr>
<td>2.13</td>
<td>Coordination Game ($\varepsilon = 0.6$) and Logit Dynamics–Basins of Attraction</td>
<td>50</td>
</tr>
<tr>
<td>2.14</td>
<td>Coordination Game and Logit Dynamics-Long-Run Welfare</td>
<td>51</td>
</tr>
<tr>
<td>2.15</td>
<td>RSP Game and Weighted Logit Dynamics</td>
<td>54</td>
</tr>
<tr>
<td>2.16</td>
<td>Schuster et. al. (1991) game and iLogit Dynamics-Period-Doubling route to chaos</td>
<td>56</td>
</tr>
<tr>
<td>2.17</td>
<td>Schuster et. al. (1991) game and iLogit Dynamics-Strange Attractor</td>
<td>57</td>
</tr>
<tr>
<td>3.1</td>
<td>Cournot duopoly with Adaptive and Rational players - Time series</td>
<td>74</td>
</tr>
</tbody>
</table>
3.2 Cournot duopoly with Adaptive and Rational players - Bifurcations, Strange Attractors .................................................. 75
3.3 Cournot duopoly game with Adaptive and Weighted Fictitious Play behaviors - Time series .................................................. 81
3.4 Cournot duopoly game with Adaptive and Weighted Fictitious Play behaviors - Bifurcations .................................................. 82
3.5 Cournot duopoly game with Adaptive and Weighted Fictitious Play behaviors - Strange Attractors ...................................... 83
3.6 Cournot duopoly game with Naive and Fictitious Play Heuristics - Time Series ................................................................. 85
3.7 Cournot duopoly game with Naive and Fictitious Play Heuristics - Attractors ................................................................. 86

4.1 Linear $n$-player Cournot game with Adaptive and Rational players .......................... 95
4.2 Linear $n$-player Cournot game with costly Rational expectations ................................ 100

5.1 Unconditional defectors (AllD) vs. Reciprocators (TFT) ........................................ 111
5.2 Reciprocators (TFT) vs Unconditional Cooperators (AllC) ..................................... 113
5.3 Unconditional defectors (AllD) vs. Unconditional cooperators (AllC) .................. 115
5.4 Unconditional defectors (AllD) vs. Generous reciprocators (GTFT) ................. 117
5.5 Unconditional defectors (AllD) vs. Pavlov (WSLS) ............................................. 119
5.6 Generous reciprocators (GTFT) vs. Unconditional cooperators (AllC) ............. 121
5.7 Generous reciprocators (GTFT) vs. Pavlovian (WSLS) ....................................... 123
5.8 Reciprocators (TFT) vs. Generous reciprocators (GTFT) .................................. 126
5.9 Reciprocators (TFT) vs. Pavlov (WSLS) ............................................................ 127
5.10 Pavlov (WSLS) vs. Unconditional cooperators (AllC) ....................................... 129
5.11 AllD vs. TFT vs. AllC ...................................................................................... 134
5.12 AllD vs. GTFT vs. WSLS - Bifurcations, Phase plots ......................................... 137
5.13 AllD vs. GTFT vs. WSLS - Bifurcation Curves .................................................. 138
5.14 AllD vs. GTFT vs. AllC ................................. 140
5.15 AllD vs. TFT vs. WSLS ................................. 143
5.16 AllD vs. TFT vs. GTFT ................................. 145
5.17 AllD vs. WSLS vs. AllC ................................. 147
5.18 TFT vs. WSLS vs. AllC ................................. 149
5.19 TFT vs. GTFT vs. WSLS ................................. 151
5.20 AllD vs. GTFT vs. WSLS vs. AllC ......................... 155
5.21 TFT vs. GTFT vs. WSLS vs. AllC ......................... 158
5.22 AllD vs. TFT vs. WSLS vs. AllC ......................... 160
5.23 AllD vs. TFT vs. GTFT vs. AllC ......................... 162
5.24 AllD vs. TFT vs. GTFT vs. WSLS ......................... 164
5.25 AllD vs. TFT vs. GTFT vs. WSLS vs. AllC - Time Series, Bifurcations 167
5.26 AllD vs. TFT vs. GTFT vs. WSLS vs. AllC - Strange Attractor . . . 168
5.27 AllD vs. TFT vs. GTFT vs. WSLS vs. AllC - Attractors . . . . . . . 169
5.28 AllD vs. TFT vs. GTFT vs. WSLS vs. AllC - Bifurcation Curves . . 171
Chapter 1

Introduction

‘Un coup de dés jamais n’abolira le hasard’ (S. Mallarmé)
(A throw of dice will never abolish chaos)

Game theory is a mathematical apparatus for dealing with decision-making in interactive, strategic environments. There are two fundamentally different views regarding the equilibria (solutions) of such strategic interactions. On the one hand, the epistemic or eductive (Binmore (1987)) approach assumes that the equilibrium is reached solely via players’ deductive reasoning about both the one-shot, interactive decision situation and other players’ reasoning. Ultimately, the decisions are to be derived from principles of rationality (Osborne and Rubinstein (1994)). On the other hand, the evolutionary approach regards the equilibrium as the steady state of some evolutionary or learning process (Hofbauer and Sigmund (2003), Weibull (1997)). Each particular strategic interaction is embedded in a sequence of past random or repeated interactions from which players aggregate information and turn it into decisions in future encounters. The distinction between the two approaches is critical for the requirements imposed on players’ rationality: the "eductive" world is inhabited by fully rational players, while the "evolutionary" world is populated by boundedly rational players. Rational players are assumed to act as expected
utility maximizers, while forming correct beliefs about the opponent play, whereas boundedly rational players face cognitive, computational and memory constraints. Hence, bounded rationality deals with:

"simple stimulus-response machines whose behavior has been tailored to their environment as a result of ill-adapted machines having been weeded out by some form of evolutionary competition"¹(Binmore (1987))

The evolutionary/learning theory of games further differentiates according to the degree of bounded rationality imposed on the interacting players. Broadly speaking, two classes of evolutionary dynamics emerge: reinforcement-based and belief-based learning models. The former class is micro-founded on reinforcing successful past, own (pure reinforcement learning) or opponent (imitation-based models) actions. At the population dynamics level it leads to, up to some possible variations, the famous Replicator Dynamics of Taylor and Jonker (1978). The belief-based class of learning models is more informationally and computationally involved because, on the one hand, it explicitly models beliefs about opponent play, and, on the other hand, it commands a better or best-response action to these beliefs. Population-wise, it yields the Fictitious Play or the Best-Reply Dynamics.

The long-run behaviour of these two classes of game dynamical processes together with its correspondence to classical game-theoretical "equilibrium" solutions (Nash equilibrium, Evolutionarily Stable Strategy, etc) has been of major concern in the evolutionary games literature. Apart from specific classes of dynamics converging to (Nash) equilibrium in certain classes of normal form games, there are still no general learning or evolutionary processes guaranteeing convergence in all classes of games. Consequently, characterizing the limiting behaviour of certain "reasonable" processes is an important area of investigation.

¹The individual rationality is effectively replaced by the ecological rationality while the restrictions imposed by the principles of rationality are replaced by restrictions on evolutionary success.
Embedded into this broad perspective, the subject of this thesis consists of the analytical and numerical study of the asymptotics of a particular learning process, the Logit Dynamics. Logit Dynamics has elements of both belief-based learning (as it can be derived from a suitable perturbation of the Best-Reply Dynamics) and reinforcement learning (as it is micro-founded in random utility theory). In essence, it predicts that players choose a best-reply to the distribution of strategies existing in the population with a probability given by the logistic function. The higher the sensitivity to payoff differences is the closer players approach the best-reply decision limit. In motivating our choice of a ‘moderately rational dynamic’ we refer back to Binmore (1987):

"Of course, the distinction between eductive and evolutive processes is quantitative rather than qualitative. In the former players are envisaged as potentially very complex machines (with very long operating costs) whereas in the latter their internal complexity is low. It is not denied that the middle ground between these extremes is more interesting than either extreme"

The Logit Dynamics captures this continuum of degrees of myopic rational behaviour: as the responsiveness to payoff differentials varies from low to high values the actual choices displayed vary from virtual random to best-reply behavior. Furthermore, the built-in behavioral parameter (intensity of choice\(^2\) in Brock and Hommes (1997)) allows systematic study of the qualitative changes (i.e. bifurcations) in the Logit Dynamics asymptotics as the behavior of players changes from random play to rational best-reply. In particular, we are interested in non-convergent, complex behavior and, more precisely, in phenomena such as path-dependence, multiple equilibria, periodic and chaotic attractors in certain appealing strategic interactions, such as:

\(^2\)Intensity of choice defines players’ sensitivity to payoffs differences between strategies. It may vary from 0 (complete insensitivity to payoff differentials, or random choice) to +∞ (highly responsive to differences in material payoffs, or best choice).
Rock-Scissors-Paper (RSP), Coordination, Cournot and Iterated Prisoner’s Dilemma games.

Before discussing in more detail the contribution of each chapter, we will briefly review the evolutionary game dynamics literature on convergence and non-convergence, both in the homogenous and heterogenous learning rules case, with a focus on periodic and chaotic limiting behaviour.

1.1 Literature on convergence of game dynamics

In lack of general convergence results, research focused on identifying certain classes of games and evolutionary dynamics, for which convergence to Nash equilibria, be it in expectation\(^3\) or in strategy, is achieved. A game is said to display the Fictitious Play Property (FPP) if every Fictitious Play process converges in expectations (beliefs). The following classes of games have been shown to have the FPP either in discrete or continuous-time\(^4\): 2-person zero-sum (Robinson (1951)), dominance solvable (Milgrom and Roberts (1991), even for more general classes of adaptive and sophisticated learning), 2-person \(2 \times 2\) nondegenerate\(^5\) (Miyasawa (1961)), 2-person \(2 \times n\) (Berger (2005)), supermodular with diminishing returns\(^6\) (Krishna (1992)), \(3 \times 3\) supermodular (Hahn (1999)), \(3 \times n\) and \(4 \times 4\) quasi-modular (Berger (2007)), weighted potential games (Monderer and Shapley (1996a)), ordinal potential games (Berger (2007)) and games of identical interest (Monderer and Shapley (1996b)). It should be noticed, that, except for the rather special classes of potential, super-

\(^3\)A discrete or continuous time process converges in expectations if the sequence of expectations approaches the set of equilibria of the game after a number of stages (Monderer and Shapley (1996b)).

\(^4\)Continuous-time Fictitious Play is equivalent, up to a time re-parametrisation, to the more frequently encountered best-response dynamics of Gilboa and Matsui (1991).

\(^5\)No player has equivalent strategies. Monderer and Sela (1996) provide an example of a degenerate \(2 \times 2\) game without FPP.

\(^6\)Games with strategic complementarities. Proof is based on a particular tie-breaking rule. Berger (2008) proves the result for ordinal strategic complementaries without relying on a specific tie-breaking rule.
modular and dominance solvable games there are no general convergence results for Fictitious Play (or best-response dynamics) in normal form games with more than 2 strategies.

The situation is somewhat similar if we turn to the second class of dynamics, namely the biologically-inspired Replicator Dynamics. Although it displays Nash stationarity\(^7\), some fixed points (even the asymptotically stable ones) may not be equilibria of the game (see the folk theorem of evolutionary game theory in Hofbauer and Sigmund (2003)). Moreover, a systematic investigation of the (continuous-time) Replicator Dynamics limit sets has only been performed up to the $3 \times 3$ games (Zeeman (1980), Bomze (1983), Bomze (1995)). The only possible attractors are saddles, sources, sinks and centers (continuum of cycles). In particular stable limit cycles (generic Hopf bifurcations) are ruled out from the asymptotic behavior of continuous-time Replicator Dynamics\(^8\) in $3 \times 3$ games.

### 1.2 Literature on complicated game dynamics

The first example of a game without the FPP was already provided by Shapley (1964) who showed, in a $3 \times 3$ bi-matrix game, that the fictitious play beliefs can cycle continuously and converge to what is known as the Shapley polygon. Richards (1997) looks at more general inductive\(^9\) learning processes in games via an alternative method to the traditional point-to-point mapping approach: she investigates mappings from regions to regions in the strategy simplex. With this region-to-region mappings, sensitive dependence on initial conditions, topological transitivity and dense periodic orbits are found for the Fictitious Play inductive rule in the ori-

---

\(^7\)All Nash equilibria of the underlying games are rest point of the dynamics.

\(^8\)Discrete-time Replicator Dynamics is already more complicated as it can display generic Hopf bifurcations (and even co-existing limit cycles) even for $3 \times 3$ games (see Weissing (1991) for an example for a non-circulant RSP payoff matrix).

\(^9\)Inductive reasoning assumes using information from the history of the play to form expectations for the future course of the game; such inductive reasoning include, besides the well-known fictitious play, Carnap dynamics, Dirichlet deliberation, etc. (Richards (1997)).
ginal Shapley bi-matrix example. Sparrow et al. (2008) and Sparrow and van Strien (2009) parametrize a Shapley $3 \times 3$ bi-matrix game and use topological arguments to show that Fictitious Play may display periodic and chaotic behavior. Building on the theory of differential inclusions they prove that, for certain parameters, the best-response dynamics displays co-existence of an attracting (but not globally attracting) Shapley polygon, chaotic orbits$^{10}$ and repelling anti-Shapley (anticlock-wise orbit) polygon. The clockwise and anticlockwise orbits exchange stability as the game payoff parameter is changed. Jordan (1993) discusses a 3-player $2 \times 2$ matching penny game and show that the FPP does not hold. Foster and Young (1998) illustrate, by example of a $6 \times 6$ coordination game, that FPP need not hold in the important class of coordination games. In their example the Fictitious Play beliefs enter a cyclical pattern. Last, Krishna and Sjöström (1998) prove a more general result, that for non-zero sum games, continuous fictitious play almost never converges cyclically to a mixed strategy equilibrium in which both players use at least three strategies.

Turning to the reinforcement learning class of dynamics, we have already observed that symmetric $3 \times 3$ games cannot yield isolated period orbits under Replicator Dynamics. The situation changes drastically when one analyses asymmetric $3 \times 3$ games or symmetric games with more than 3 strategies. For instance, Aguiar and Castro (2008) investigate the asymptotic behavior of reinforcement learning (or replicator dynamics) in a $3 \times 3$ RSP bi-matrix game and provide geometric proofs for chaotic switching near the heteroclinic network formed by the 9 vertices rest points of the bi-matrix replicator dynamics. Their results confirm the chaotic attractors discovered numerically in the bimatrix RSP game (Sato et al. (2002)). In $4 \times 4$ symmetric games, limit cycles with Replicator Dynamics are reported in, for instance, Hofbauer (1981), Akin (1982), Maynard Smith and Hofbauer (1987). Limit cycles and irrational behavior$^{11}$ are exhibited in a $4 \times 4$ symmetric game constructed by Berger and Hofbauer

---

$^{10}$ Of "jitter type", i.e. neither attracting, repelling or of saddle type

$^{11}$ Survival of strictly dominated strategies.

So far, the implicit assumption was that players are homogenous with respect to the learning heuristic used (be it of fictitious play, reinforcement, or imitation type) when beliefs are formed about the opponent’s play. However, this need not be the case in, for instance, multi-population games when each population is endowed with a particular expectation-formation rule, or, when players are allowed to switch their learning procedures from one game interaction to the other. Kaniovski et al. (2000)\textsuperscript{12} study 2-population $2 \times 2$ coordination and anti-coordination games with players in each population falling into one of the following three categories (best-responders, imitators/conformists and anti-conformists/contrarians). The fractions are fixed and there is no opportunity to switch between learning rules, i.e. there is no room for ‘learning how to learn’. Using the Poincare-Bendixon theorem, they find, for certain initial mixtures of the three categories of heuristics, convergence to a limit cycle for these 2-player generic $2 \times 2$ games.

1.3 Thesis Outline

This thesis is structured in four self-contained chapters, each with its own introduction, conclusions and appendices, which can be read independently. The bibliography is collected at the end of the thesis. Although each chapter investigates different games, they are nevertheless related through the Logit evolutionary dynamic put at work and the general techniques from bifurcation theory in nonlinear dynamical

\textsuperscript{12}To our knowledge, this is one of the few papers dealing with heterogeneous learning rules in deterministic game dynamics. For evolutionary competition between a best-response and an imitation rule, in a $2\times2$ coordination game, mutation-selection setting a la Kandori et al. (1993), we refer the reader to Juang (2002).
systems employed. We find Logit Dynamics an appealing and reasonable way of modeling bounded rationality given that it shares elements of both reinforcement and belief-learning. At the same time, the two extreme cases of random and rational play are special cases of the Logit Dynamics specification. Furthermore, it can also be used to model heterogenous learning, i.e. switching between alternative heuristics.

Chapter 2 investigates the behaviour of simple RSP and $3 \times 3$ Coordination games when all players are homogenous, i.e. they employ the same logistic updating mechanism when given the opportunity to revise their status-quo strategy. Heterogenous learning rules and evolutionary competition between different heuristics is introduced in Chapters 3 and 4 in the context of linear 2 and $n$-player Cournot games. In Chapter 5 we discuss the evolution of Iterated Prisoner’s Dilemma meta-rules where players update their repeated strategies according to the same logistic protocol.

1.3.1 Multiple Steady States, Limit Cycles and Chaotic Attractors in Logit Dynamics

The starting point of Chapter 2 are two results proved in Zeeman (1980). First, all Hopf bifurcations are degenerate in 2-person $3 \times 3$ stable symmetric games under (continuous-time) Replicator Dynamics. Thus, no isolated periodic orbits are possible in RSP symmetric games and the limiting behavior coincides either with a continuum of cycles or a heteroclinic cycle connecting the three monomorphic steady states. Second, Replicator Dynamics on a $3 \times 3$ stable game can have at most one, interior fixed point within the 2-simplex. In this chapter we study whether these two results hold also for a ‘rationalistic’ dynamics, the smooth version of best-reply dynamics, the Logit Dynamics, in the context of two interesting classes of $3 \times 3$ symmetric games: Rock-Paper-Scissors and Coordination game. For the circulant$^{14}$ RSP

---

$^{13}$Zeeman (1980) uses this notion in the sense of structural stability: "a game is stable if sufficiently small perturbations of its payoff matrix induce topologically equivalent flows".

$^{14}$The normal form payoff matrix is circulant.
game we first give an alternative proof to Zeeman’s result, based on the computation of the Lyapunov coefficient in the normal form of the vector field and show that for the replicator dynamics the higher order terms degeneracy condition is always satisfied. This rules out the possibility of a stable limit cycle through a generic Hopf bifurcation. Second, via the same technique we show that there are generic Hopf bifurcations in the Logit Dynamics on these $3 \times 3$ circulant payoff matrix RSP games. Moreover, the Hopf bifurcation is always supercritical i.e. the interior, isolated limit cycle is born stable and attracts the entire two-dimensional simplex. This result is extended, through numerical and continuation analysis using the bifurcation software Matcont\textsuperscript{15} (Dhooge et al. (2003)) to the class of non-circulant symmetric RSP games.

As far as the second result in Zeeman (1980) is concerned, in a $3 \times 3$ Coordination game we detect, numerically, a route to multiple interior steady states of the Logit Dynamics, via a sequence of three fold bifurcations. The basins of attraction of the three steady states vary non-monotonically with respect to the behavioral parameter (the sensitivity to payoff differences) with maximal welfare\textsuperscript{16} attained only for moderate values of rationality.

All detected singularities—supercritical Hopf and fold—are then "continued" - i.e. followed in the parameter space using the Matcont software - both in the game payoff matrix parameters and logistic choice behavioral parameter. The bifurcation curves are found to be robust to perturbations in all these parameters.

Furthermore, a frequency-weighted version of the Logit Dynamics is run on a circulant RSP game and found to inherit dynamic/long-run features from both Replicator and Logit dynamics. Weighted Logit exhibits supercritical Hopf bifurcations leading to stable limit cycles reminiscent of Logit dynamics, but of larger amplitude than the limit cycles in the Logit Dynamics. These limit cycles approach the hetero-

\textsuperscript{15}A continuation software for ode (see package documentation at http://www.matcont.ugent.be/)

\textsuperscript{16}The measure of welfare is constructed as the payoff at steady state weighted by the size of the corresponding basins of attraction.
clinic cycle of Replicator Dynamics as the intensity of choice vanishes. Finally, the same weighted version of the Logit transforms the periodic dynamics of Replicator Dynamics into chaotic fluctuations on a $4 \times 4$ symmetric game matrix inspired from the biological literature.

### 1.3.2 Heterogenous Learning Rules in Cournot Games

It is an established results that homogenous expectations of fictitious play type are conducive to Cournot-Nash equilibrium (henceforth, CNE) play in 2 (Deschamp (1975)) and $n$-player (Thorlund-Petersen (1990)) linear demand-linear cost Cournot games. These results are extended to general adaptive and sophisticated learning processes by Milgrom and Roberts (1991). For linear demand-quadratic costs, discrete strategy set duopoly, Cox and Walker (1998) show that best-reply or Fictitious Play still converges to CNE as long as the marginal costs are not decreasing too fast\(^{17}\) (a situation dubbed Type I duopoly by Cox and Walker (1998)). However, when one or both firms’ marginal cost decrease rapidly enough\(^{18}\) (a game coined Type II duopoly), Cox and Walker (1998) prove that with homogenous Cournot (i.e. naive) expectations, the interior CNE loses stability and, depending on initial conditions, the system may converge to either one of the two boundary NE or to a 2-cycle of the form \{$(0,0), (q^{BNE}, q^{BNE}), 0 < q^{CNE} < q^{BNE}$\}

Chapter 3 builds an evolutionary version of such a type II, linear demand-quadratic costs Cournot duopoly with heterogenous players and heuristic switching, similar as in Droste et al. (2002). We enrich their ecology with a long memory weighted fictitious play rule and study both analytically and numerically the log-run behaviour of logistic switching between predictors in various $2 \times 2$ rules sub-ecologies. Our focus is on the interplay between different learning heuristics and on the impact

---

\(^{17}\)The two firms’ reaction function intersect in a normal way, i.e. they have only one (interior) intersection.

\(^{18}\)Such that the reaction curves have two additional "boundary" intersections besides the interior CNE.
of the evolutionary competition between expectation-formation rules on the stability of the Cournot-Nash equilibrium. The heuristics toolbox consists of naive, adaptive, fictitious and weighted fictitious play and rational (Nash) expectations. It is also assumed, because of more demanding information gathering effort, that the more sophisticated predictors in each $2 \times 2$ subecology are costly. This creates incentives, as in Brock and Hommes (1997) for players to switch away from complicated predictors when the system lingers in a stable regime near the interior CNE and back into the costly sophisticated predictor in a "turbulent" far from CNE regime. We first generalize an ecology of naive and Nash expectation in Droste et al. (2002) to allow for adaptive expectations, derive the corresponding instability thresholds and find qualitatively similar long-run behavior. Second, an ecology of endogenously selected, free adaptive and costly weighted fictitious play expectations is shown to destabilize the unique interior CNE - as, for instance, the intensity of choice parameter increases—through a period-doubling bifurcation followed by a secondary Neimark-Sacker singularity and, eventually, collapsing into chaos.

1.3.3 On the Stability of the Cournot Solution: An Evolutionary Approach

Chapter 4 reverts to the simplest possible linear-cost linear demand Cournot framework and investigates the stability of the CNE for an arbitrary number of players. It revisits, in an heterogenous players, evolutionary set-up, the famous Theocharis (1960) instability result: in a linear Cournot game, the CNE is neutrally stable under naive expectations for 3 players and unstable, with unbounded oscillations, for more than 3 players. Players are randomly matched to play a $n$-person quantity-setting game and, similar to the previous chapter, they may choose between a costless adaptive expectations predictor and a costly more involved rational predictor. The CNE number of players instability threshold is derived as a function of the degree of expectations adaptiveness, costs of rational expectations and sensitivity to heurist-
ics’ differential performance. In a heterogenous agents setting with agents switching between adaptive and rational heuristics Theocharis (1960) result is re-evaluated: as the number of players increases the system destabilizes through a period-doubling route to chaos. Theocharis (1960) triopoly instability result obtains as the limit of a system with naive and homogenous (no switching) expectation. In the quadropoly case, the unstable CNE can now be "stabilized" by fine-tuning model parameters; for instance, by making the expectations more adaptive (i.e. looking more into the past).

1.3.4 Evolution in Iterated Prisoner’s Dilemma Games under Smoothed Best-Reply Dynamics

Emergence and evolution of cooperation in social dilemmas rank among the most salient departures from the predictions of rational actor (game) theory. The repairs in the spirit of "folk" theorems of repeated games are far from satisfactory, as long as they rely on unrealistically heavy discounting of future payoffs. The boundedly rationality models with heterogenous players endowed with "smart and simple" heuristics (Gigerenzer and Todd (1999)) emerged as a promising alternative for understanding the puzzles of ubiquitous cooperation observed in human societies.

An ecology of such simple heuristics for dealing with repeated Prisoner's Dilemma interaction is constructed in Chapter 5. Its evolution and limiting behavior are investigated as players receive opportunities to revise status-quo and update to a better-performing meta-rule available in the population.

The model builds an ecology of "tit-for-tat" reciprocators ($TFT$), undiscriminating defectors ($AllD$), and cooperators ($AllC$) as in Sigmund and Brandt (2006), extended with the Pavlovian "win-stay-lose-shift" ($WSLS$) heuristic\footnote{Pavlov, or stimuli-driven players, preserve the status-quo as long as they are satisfied with the achieved performance and proceed to exploring alternatives otherwise. Hence, the win-stay-lose-shift $WSLS$ label.} and the gen-
erous variation of homo reciprocator "generous tit-for-tat" ($GTFT$)\textsuperscript{20}. The resulting subecologies of $2 \times 2$, $3 \times 3$, $4 \times 4$ together with the complete five rules ecology are discussed analytically and numerically with rich dynamics unfolding as the heuristics’ toolbox enlarges. An abundance of Rock-Paper-Scissors like patterns is discovered in the $3 \times 3$ ecologies comprising Pavlovian and "generous" players, while some $4 \times 4$ ecologies display path-dependence along with co-existence of periodic and chaotic attractors. Turning to the performance of our heuristics selection, the surrounding ecology appears critical for the success or failure of a particular repeated game meta-rule. For instance, the stimulus-response strategy does well in a no $AllC$ $4 \times 4$ environment but poorly when unconditional cooperators are around. However, in the full $5 \times 5$ ecology Pavlov almost goes extinct in the best-reply limit of the logit dynamics, but spreads to a large fraction of the population when players are boundedly rational.

\textsuperscript{20}$GTFT$ interprets opponent defection as mistakes and, with certain positive probability, resets cooperation.
Chapter 2

Multiple Steady States, Limit Cycles and Chaotic Attractors in Logit Dynamics

2.1 Introduction

2.1.1 Motivation

A large part of the research on evolutionary game dynamics focused on identifying conditions that ensure uniqueness of and/or convergence to point-attractors such as Nash Equilibrium and Evolutionary Stable Strategy (ESS). Roughly speaking, within this ‘convergence’ literature one can further distinguish between literature focusing on classes of games (e.g. Milgrom and Roberts (1991), Nachbar (1990), Hofbauer and Sandholm (2002)) and literature on the classes of evolutionary dynamics (e.g. Cressman (1997), Hofbauer (2000), Hofbauer and Weibull (1996), Sandholm (2005), Samuelson and Zhang (1992)). Nevertheless, there are some examples of periodic and chaotic behaviour in the literature, mostly under a particular kind of evolutionary dynamics, the Replicator Dynamics. Motivated by the idea of adding an explicit
dynamical process to the static concept of Evolutionary Stable Strategy, Taylor and Jonker (1978) introduced the Replicator Dynamics. It soon found applications in biological, genetic or chemical systems, those domains where organisms, genes or molecules evolve over time via replication. The common feature of these systems is that they can be well approximated by an infinite population game with random pairwise matching, giving rise to a replicator-like evolutionary dynamics given by a low-dimensional dynamical system.

In the realm of the non-convergence literature two issues are particularly important: the existence of stable periodic and complicated solutions and their robustness(to slight perturbations in the payoffs matrix). Hofbauer et al. (1980) and Zeeman (1981) investigate the phase portraits resulting from three-strategy games under the replicator dynamics and conclude that only ‘simple’ behaviour - sinks, sources, centers, saddles - can occur. In general, an evolutionary dynamics together with a $n-$strategy game define a proper $n − 1$ dynamical system on the $n − 1$ simplex. An important result the proof of Zeeman (1980) that there are no generic Hopf bifurcations on the 2-simplex: "When $n = 3$ all Hopf bifurcations are degenerate". Bomze (1983), Bomze (1995) provide a thorough classification of the planar phase portraits for all 3-strategy games according to the number, location(interior or boundary of the simplex) and stability properties of the Replicator Dynamics fixed points and identify 49 different phase portraits: again, only non-robust cycles are created usually via a degenerate Hopf bifurcation.

Hofbauer (1981) proves that in a 4-strategy game stable limit cycles are possible under Replicator Dynamics; the proof consists in finding a suitable Lyapunov function whose time derivative vanishes on the $\omega-$limit set of a periodic orbit. Stable limit cycles are also reported in Akin (1982) in a genetic model where gene ‘replicates’ via the two allele-two locus selection; this is not surprising as the dynamical system modeling gametic frequencies is three-dimensional, the dynamics is of Replicator type.

---

1Zeeman (1980), pp. 493
and the Hofbauer (1981) proof applies here, as well. Furthermore, Maynard Smith and Hofbauer (1987) prove the existence of a stable limit cycle for an asymmetric, Battle of Sexes-type genetic model where the allelic frequencies evolution defines again a 3-D system. Their proof hinges on normal form reduction together with averaging and elliptic integrals techniques for computing the phase and angular velocity of the periodic orbit. Stadler and Schuster (1990) perform an impressive systematic search for both generic (fixed points exchanging stability) and degenerate (stable and unstable fixed points colliding into a one or two-dimensional manifold at the critical parameter value) transitions between phase portraits of the replicator equation on $3 \times 3$ normal form game.

Chaotic behaviour is found by Schuster et al. (1991) in Replicator Dynamics for a 4-strategy game matrix derived from an autocatalytic reaction network. They report the standard Feigenbaum route to chaos: a cascade of period-doubling bifurcations intermingled with several interior crisis and collapses to a chaotic attractor. Numerical evidence for strange attractors is provided by Skyrms (1992), Skyrms (2000) for a Replicator Dynamics flow on two examples of a four-strategy game.

Although periodic and chaotic behaviour is substantially documented in the literature for the Replicator Dynamics, there is much less evidence for such complicated behaviour in classes of evolutionary dynamics that are more appropriate for humans interaction (fictitious play, best response dynamics, adaptive dynamics, etc.). While Shapley (1964) constructs an example of a non-zero sum game with a limit cycle under fictitious play and Berger and Hofbauer (2006) find stable periodic behaviour - two limit cycles bounding an asymptotically stable annulus - for a different dynamic - the Brown-von Newmann Nash (BNN) - a systematic characterization of (non) generic bifurcations of phase portraits is still missing for the Best Response dynamics.

In this Chapter we take a first step in this direction and use a smoothed version of the Best Response dynamics - the Logit Dynamics - to study evolutionary dynamics
in simple three and four strategies games from the existing literature. The qualitative behaviour of the resulting ‘evolutionary’ games is investigated with respect to changes in the payoff and behavioural parameters, using analytical and numerical tools from non-linear dynamical system theory.

2.1.2 ‘Replicative’ vs. ‘rationalistic’ dynamics

Most of the earlier discussed evolutionary examples are inspired from biology and not from social sciences. They concerned animal contests (Zeeman (1980), Bomze (1983), Bomze (1995)), genetics (Maynard Smith and Hofbauer (1987)) or chemical catalytic networks (Schuster et al. (1991), Stadler and Schuster (1990)). From the perspective of strategic interaction the main criticism of the ‘biological’ game-theoretic models is targeted at the intensive use of preprogrammed, simple imitative play with no role for optimization and innovation. Specifically, in the transition from animal contests and biology to humans interactions and economics the Replicator Dynamics seems no longer adequate to model the rationalistic and ‘competent’ forms of behaviour (Sandholm (2008)). Best Response Dynamics would be more applicable to human interaction as it assumes that agents are able to optimally compute and play a (myopic) ‘best response’ to the current state of the population. But, while the Replicator Dynamics appeared to impose an unnecessarily loose rationality assumption the Best Response dynamics moves to the other extreme: it is too stringent in terms of rationality. Another drawback is that, technically, the best reply is not necessarily unique and this leads to a differential inclusions instead of an ordinary differential equation. One way of solving these problems was to stochastically perturb the matrix payoffs and derive, via the discrete choice theory, a ‘noisy’ Best Response Dynamics, called the Logit Dynamics. Mathematically it is a ‘smoothed’, well-behaved dynamics while from the strategic interaction point of view it models a boundedly rational player/agent. Moreover, from a nonlinear dynamical systems perspective the Replicator Dynamic is non-generic in dimension two and only degenerate Hopf bifurcations
can arise on the 2-simplex. In sum, apart from its conjectured generic properties, the Logit is recommended by the need for modelling players with different degrees of rationality and for smoothing the Best Response correspondence.

Thus, the main part of this Chapter investigates the qualitative behaviour of simple evolutionary games under the Logit dynamics when the level of noise and/or underlying normal form game payoffs matrix is varied with the goal of finding attractors (periodic or perhaps more complicated) which are not fixed points or Nash Equilibria.

This Chapter is organized as follows: Section 2.2 introduces the Logit Dynamics, while Section 2.3 gives a brief overview of the Hopf bifurcation theory. In Sections 2.4 the Logit Dynamics is implemented on various versions of Rock-Scissors-Paper and Coordination games while Section 2.5 discusses an example of chaotic dynamics under a frequency-weighted version of the Logit Dynamics. Section 2.6 contains concluding remarks.

2.2 The Logit Dynamics

2.2.1 Evolutionary dynamics

Evolutionary game theory deals with games played within a (large) population over a long time horizon (evolution scale). Its main ingredients are the underlying normal form game - with payoff matrix $A[n \times n]$ - and the evolutionary dynamic class which defines a dynamical system on the state of the population. In a symmetric framework, the strategic interaction takes the form of random matching with each of the two players choosing from a finite set of available strategies $E = \{E_1, E_2,...E_n\}$. For every time $t$, $x(t)$ denotes the $n$-dimensional vector of frequencies for each strategy/type $E_i$ and belongs to the $n - 1$ dimensional simplex $\Delta^{n-1} = \{x \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 1\}$.

Under the assumption of random interactions strategy $E_i$ fitness would be simply determined by averaging the payoffs from each strategic interaction with weights
given by the state of the population $x$. Denoting with $f(x)$ the payoff vector, its components - individual payoff or fitness of strategy $i$ in biological terms - are:

$$f_i(x) = (Ax)_i$$  \hfill (2.1)

Sandholm (2006) rigorously defines an *evolutionary dynamics* as a map assigning to each population game a differential equation $\dot{x} = V(x)$ on the simplex $\Delta^{n-1}$. In order to derive such an 'aggregate' level vector field from *individual choices* he introduces a *revision protocol* $\rho_{ij}(f(x), x)$ indicating, for each pair $(i, j)$, the rate of switching($\rho_{ij}$) from the currently played strategy $i$ to strategy $j$. The mean vector field is obtained as:

$$\dot{x}_i = V_i(x) = \text{inflow into strategy } i - \text{outflow from strategy } i$$

$$= \sum_{j=1}^{n} x_j \rho_{ji}(f(x), x) - x_i \sum_{j=1}^{n} \rho_{ij}(f(x), x).$$ \hfill (2.2)

Based on the computational requirements/quality of the revision protocol $\rho$ the set of evolutionary dynamics splits into two large classes: imitative dynamics and pairwise comparison (‘competent’ play). The first class is represented by the most famous dynamics, the Replicator Dynamic (Taylor and Jonker (1978)) which can be easily derived by substituting into (2.2) the pairwise proportional revision protocol:

$$\rho_{ij}(f(x), x) = x_j [f_j(x) - f_i(x)]_+ \text{ (player } i \text{ switches to strategy } j \text{ at a rate proportional with the probability of meeting an } j\text{-strategist}(x_j) \text{ and with the excess payoff of opponent } j-[f_j(x) - f_i(x)]\text{ iff positive):}$$

$$\dot{x}_i = x_i[f_i(x) - \bar{f}(x)] = x_i[(Ax)_i - xAx]$$ \hfill (2.3)

where $\bar{f}(x) = xAx$ is the average population payoff.

Although widely applicable to biological/chemical models, the Replicator Dynamics lacks the proper individual choice, micro-foundations which would make it
attractive for modelling humans interactions. The alternative - Best Response dynamic - already introduced by Gilboa and Matsui (1991) requires extra computational abilities from agents, beyond merely sampling randomly a player and observing the difference in payoff: specifically being able to compute a best reply strategy to the current population state:

\[ \dot{x}_i = BR(x) - x_i \]  
where,

\[ BR(x) = \arg \max_y yf(x) \]

### 2.2.2 Discrete choice models-the Logit choice rule

Apart from the highly unrealistic assumptions regarding agents capacity to compute a perfect best reply to a given population state there is also the drawback that (2.4) defines a differential inclusion, i.e. a set-valued function. The best responses may not be unique and multiple trajectory paths can emerge from the same initial conditions. A ‘smoothed’ approximation of the Best Reply dynamics - the Logit dynamics - was introduced by Fudenberg and Levine (1998); it was obtained by stochastically perturbing the payoff vector \( f(x) \) and deriving the Logit revision protocol:

\[ \rho_{ij}(f(x), x) = \exp[\eta^{-1}f_j(x)] / \sum_k \exp[\eta^{-1}f_k(x)] = \exp[\eta^{-1}Ax]_i / \sum_k \exp[\eta^{-1}Ax]_k, \]  

where \( \eta > 0 \) is the noise level. Here \( \rho_{ij} \) represents the probability of player \( i \) switching to strategy \( j \) when provided with a revision opportunity. For high levels of noise the choice is fully random (no optimization) while for \( \eta \) close to zero the switching probability is almost one. This revision protocol can be explicitly derived from a random utility model or discrete choice theory (McFadden (1981), Anderson et al. (1992)) by adding to the payoff vector \( f(x) \) a noise vector \( \varepsilon \) with a particular distribution, i.e. \( \varepsilon_i \) are i.i.d following the extreme value distribution \( G(\varepsilon_i) = \exp(-\exp(-\eta^{-1}\varepsilon_i - \gamma)), \gamma = 0.5772 \) (the Euler constant). The density of
this Weibull type distribution is

$$g(\varepsilon_i) = G'(\varepsilon_i) = \eta^{-1} \exp(-\eta^{-1}\varepsilon_i - \gamma) \exp(-\exp(-\eta^{-1}\varepsilon_i - \gamma)).$$

With noisy payoffs, the probability that strategy $E_i$ is a best response can be computed as follows:

$$P(i = \arg \max_j [(Ax)_i + \varepsilon_i] = P([A(x)_i + \varepsilon_i > (Ax)_j + \varepsilon_j], \forall j \neq i$$

$$= P[\varepsilon_j < (Ax)_i + \varepsilon_i - (Ax)_j], j \neq i = \int_{-\infty}^{\infty} g(\varepsilon_i) \prod_{j \neq i} G((Ax)_i + \varepsilon_i - (Ax)_j) d\varepsilon_i$$

$$= \int_{-\infty}^{\infty} \eta^{-1} \exp(-\eta^{-1}\varepsilon_i - \gamma) e^{-\exp(-\eta^{-1}\varepsilon_i)} \prod_{j \neq i} e^{-\exp(-\eta^{-1}[(Ax)_i + \varepsilon_i - (Ax)_j] - \gamma)} d\varepsilon_i$$

which simplifies to our logit probability of revision:

$$\rho_{ij} = \frac{\exp[\eta^{-1}(Ax)_i]}{\sum_k \exp[\eta^{-1}(Ax)_k]}$$

An alternative way to obtain (2.5) is to deterministically perturb the set-valued best reply correspondence (2.4) with a strictly concave function $V(y)$ (Hofbauer (2000)):

$$BR_{\eta}(x) = \arg \max_{y \in \Delta^{n-1}} [y \cdot (Ax) + V_{\eta}(y)]$$

For a particular choice of the perturbation function $V_{\eta}(y) = \eta \sum_{i=1}^{n} y_i \log y_i, y \in \Delta^{n-1}$ the resulting objective function is single-valued and smooth; the first order condition yields the unique logit choice rule:

$$BR_{\eta}(x)_i = \frac{\exp[\eta^{-1}(Ax)_i]}{\sum_k \exp[\eta^{-1}(Ax)_k]}$$

Plugging the Logit revision protocol (2.5) back into the general form of the mean field dynamic (2.2) and making the substitution $\beta = \eta^{-1}$ we obtain a well-behaved
system of o.d.e.’s, the Logit dynamics as a function of the intensity of choice (Brock and Hommes (1997)) parameter $\beta$:

$$
\dot{x}_i = \frac{\exp[\beta A x_i]}{\sum_k \exp[\beta A x_k]} - x_i
$$

When $\beta \to \infty$ the probability of switching to the discrete ‘best response’ $j$ is close to one while for a very low intensity of choice ($\beta \to 0$) the switching rate is independent of the actual performance of the alternative strategies (equal probability mass is put on each of them). The Logit dynamics displays some properties characteristic to the logistic growth function, namely high growth rates($\dot{x}_i$) for small values of $x_i$ and growth ‘levelling off’ when close to the upper bound. This means that a specific frequency $x_i$ grows faster when it is already large in the Replicator Dynamics relative to the Logit dynamics.

### 2.3 Hopf and degenerate Hopf bifurcations

As the main focus of the thesis is the detection of stable cyclic behaviour this section will shortly review the main bifurcation route towards periodicity, the Hopf bifurcation. In a one-parameter family of continuous-time systems, the only generic bifurcation through which a limit cycle is created or disappears is the non-degenerate Hopf bifurcation. The planar case will be discussed first and then, briefly, the methods to reduce higher-dimensional systems to the two-dimensional one. The main mathematical result (see, for example, Kuznetsov (1995)) is:

Assume we are given a parameter-dependent, two dimensional system:

$$
x = f(x, \alpha), x \in \mathbb{R}^2, \alpha \in \mathbb{R}, \text{smooth}
$$

with the Jacobian matrix evaluated at the fixed point $x^* = 0$ having a pair of purely
imaginary, complex conjugates eigenvalues:

\[ \lambda_{1,2} = \mu(\alpha) \pm i\omega(\alpha), \mu(\alpha) < 0, \alpha < 0, \mu(0) = 0 \]

and \( \mu(\alpha) > 0, \alpha > 0 \).

If, in addition, the following genericity\(^2\) conditions are satisfied:

(i) \( \left[ \frac{\partial \mu(\alpha)}{\partial \alpha} \right]_{\alpha=0} \neq 0 - \text{transversality condition} \)

(ii) \( l_1(0) \neq 0 \), where \( l_1(0) \) is the first Lyapunov coefficient\(^3\) - nondegeneracy condition,

then the system (2.7) undergoes a Hopf bifurcation at \( \alpha = 0 \). As \( \alpha \) increases the steady state changes stability from a stable focus into an unstable focus.

There are two types of Hopf bifurcation, depending on the sign of the first Lyapunov coefficient \( l_1(0) \):

(a) If \( l_1(0) < 0 \) then the Hopf bifurcation is supercritical: the stable focus \( x \) becomes unstable for \( \alpha > 0 \) and is surrounded by an isolated, stable closed orbit (limit cycle).

(b) If \( l_1(0) > 0 \) then the Hopf bifurcation is subcritical: for \( \alpha < 0 \) the basin of attraction of the stable focus \( x^* \) is surrounded by an unstable cycle which shrinks and disappears as \( \alpha \) crosses the critical value \( \alpha = 0 \) while the system diverges quickly from the neighbourhood of \( x^* \).

In the first case the stable cycle is created immediately after \( \alpha \) reaches the critical value and thus the Hopf bifurcation is called supercritical, while in the latter the unstable cycle already exists before the critical value, i.e. a subcritical Hopf bifurcation (Kuznetsov (1995)). The supercritical Hopf is also known as a soft or

\(^2\)Genericity usually refers to transversality and non-degeneracy conditions. Roughly speaking, the transversality condition means that eigenvalues cross the real line at non-zero speed. The nondegeneracy condition implies non-zero higher-order coefficients in equation (2.10) below. It ensures that the singularity \( x^* \) is typical (i.e. ‘nondegenerate’) for a class of singularities satisfying certain bifurcation conditions (see Kuznetsov (1995)).

\(^3\)This is the coefficient of the third order term in the normal form of the Hopf bifurcation (see equation (2.10) below).
non-catastrophic bifurcation because, even when the system becomes unstable, it still lingers within a small neighbourhood of the equilibrium bounded by the limit cycle, while the subcritical case is a sharp/catastrophic one as the system now moves quickly far away from the unstable equilibrium.

If $l_1(0) = 0$ then there is a degeneracy in the third order terms of normal form and, if other, higher order nondegeneracy conditions hold (i.e. non-vanishing second Lyapunov coefficient) then the bifurcation is called Bautin or generalized Hopf bifurcation. This happens when the first Lyapunov coefficient vanishes at the given equilibrium $x^*$ but the following higher-order genericity conditions hold:

(i) $l_2(0) \neq 0$, where $l_2(0)$ is the second Lyapunov coefficient - nondegeneracy condition

(ii) the map $\alpha \to (\mu(\alpha), l_1(\alpha))$ is regular (i.e. the Jacobian matrix is nonsingular) at the critical value $\alpha = 0$ - transversality condition.

Depending on the sign of $l_2(0)$, at the Bautin point the system may display a limit cycle bifurcating into two or more cycles, coexistence of stable and unstable cycles which collide and disappear, together with cycle blow-up.

**Computation of the first Lyapunov coefficient**

For the planar case, $l_1(0)$ can be computed without explicitly deriving the normal form, from the Taylor coefficients of a transformed version of the original vector field. The computation of $l_1(0)$ for higher dimensional systems makes use of the Center Manifold Theorem by which the orbit structure of the original system near $(x^*, 0)$ is fully determined by its restriction to the two-dimensional center manifold. On

---

4Technically, these ‘higher-order’ genericity conditions ensure that there are smooth invertible coordinate transformations, depending smoothly on parameters, together with parameter changes and (possibly) time re-parametrizations such that (2.7) can be reduced to a ‘simplest’ form, the normal form. See Kuznetsov (1995) pp. 309 for more details on the Bautin (generalized Hopf) bifurcation and for an expression for the second Lyapunov coefficient $l_2(0)$.

5The center manifold is the manifold spanned by the eigenvectors corresponding to the eigenvalues with zero real part.
the center manifold (2.7) takes the form (Wiggins (2003)): 

\[
\begin{pmatrix}
\dot{x} \\
\dot{y}
\end{pmatrix} = \begin{pmatrix}
\text{Re}\lambda(\delta) & -\text{Im}\lambda(\delta) \\
\text{Im}\lambda(\delta) & \text{Re}\lambda(\delta)
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix} + \begin{pmatrix}
f^1(x, y, \delta) \\
f^2(x, y, \delta)
\end{pmatrix}
\] (2.8)

where \(\lambda(\delta)\) is an eigenvalue of the linearized vector field around the steady state and \(f_1(x, y, \delta), f_2(x, y, \delta)\) are nonlinear functions of order \(O(|x|^2)\) to be obtained from the original vector field. Wiggins (2003) also provides a procedure for transforming (2.7) into (2.8). Specifically, for any vector field \(\dot{x} = F(x), x \in \mathbb{R}^2\) let \(DF(x^*)\) denote the Jacobian evaluated at the fixed point \(x^*\). Then \(\dot{x} = F(x)\) is equivalent to:

\[
\dot{x} = Jx + T^{-1}F(Tx)
\] (2.9)

with \(J\) stands for the real Jordan canonical form of \(DF(x^*)\), \(T\) is the matrix transforming \(DF(x^*)\) into the Jordan form, and \(\bar{F}(x) = F(x) - DF(x^*)x\). At the Hopf bifurcation point \(\delta_0, \lambda_{1,2} = \pm i\omega\) and the first Lyapunov coefficient is (Wiggins (2003)):

\[
l_1(\delta) = \frac{1}{16} \left[ f^1_{xxx} + f^1_{xyy} + f^2_{xxy} + f^2_{yy} \right] + \\
+ \frac{1}{16\omega} \left( f^1_{xy} (f^1_{xx} + f^1_{yy}) - f^2_{xy} (f^2_{xx} + f^2_{yy}) - f^1_{xx} f^2_{xx} + f^1_{yy} f^2_{yy} \right)
\] (2.10)

\[
(2.11)
\]

### 2.4 Three strategy games

We consider two well-known three-strategy games: the generalized Rock-Scissors-Paper and the Coordination Game. In subsection (2.4.1) we first perform a local bifurcation analysis for a classical example of three-strategy games, the Rock-Scissors-Paper. Two types of evolutionary dynamics - Replicator and Logit Dynamics - are considered while the qualitative change in their orbital structure is studied with respect to the behavioural parameter \((\beta)\) and the payoff \((\varepsilon, \delta)\) parameters. In subsection (2.4.2) we provide numerical evidence for a sequence of fold bifurcations in
the Coordination game under the Logit Dynamics and depict the fold curves in the parameter space.

2.4.1 Rock-Scissors-Paper Games

The Rock-Paper-Scissors class of games (or games of cyclical dominance) formalize strategic interactions where each strategy $E_i$ is an unique best response to strategy $E_{i+1}$ for $i = 1, 2$ and $E_3$ is a best response to $E_1$.

$$A = \begin{pmatrix} \gamma_1 & \delta_2 & \varepsilon_3 \\ \varepsilon_1 & \gamma_2 & \delta_3 \\ \delta_1 & \varepsilon_2 & \gamma_3 \end{pmatrix}; \delta_i \geq \gamma_i \geq \varepsilon_i \tag{2.12}$$

Due to the invariance under positive linear transformations of the payoff matrix (2.12) (Zeeman (1980), Weissing (1991)) the main diagonal element can be set to zero, by, for instance, substracting the diagonal entry from each column entries:

$$A = \begin{pmatrix} 0 & \delta_2 & -\varepsilon_3 \\ -\varepsilon_1 & 0 & \delta_3 \\ \delta_1 & -\varepsilon_2 & 0 \end{pmatrix}; \delta_i, \varepsilon_i \geq 0 \tag{2.13}$$

If matrix (2.13) is circulant (i.e. $\delta_i = \delta, \varepsilon_i = \varepsilon, i = 1, 2, 3$) then the RSP game is called circulant, while for a non-circulant matrix (2.13) we have a generalized RSP game. The behavior of Replicator and Logit Dynamics on the class of circulant RSP games will be investigated, both analytically and numerically, in the first and second part of this subsection, respectively. Numerical results about the generalized RSP class of games under Logit Dynamics are reported in the third part.

---

6 see Weissing (1991) for a thorough characterization of the discrete-time Replicator Dynamics behavior on this class of games.

7 We stick to the same notations as in (2.12) although it is apparent that the new $\delta'$s and $\varepsilon'$s are different from the old ones as they are derived via the above mentioned linear transformation.
Circulant RSP Game and Replicator Dynamics

The circulant RSP game is a first generalization of the classical, zero-sum form of RSP game as discussed in, for instance, Hofbauer and Sigmund (2003):

\[
A = \begin{pmatrix}
0 & \delta & -\varepsilon \\
-\varepsilon & 0 & \delta \\
\delta & -\varepsilon & 0 \\
\end{pmatrix}, \delta, \varepsilon > 0
\] (2.14)

Letting \( \mathbf{x}(t) = (x(t), y(t), z(t)) \) denote the population state at time instance \( t \) define a point from the 2-dimensional simplex, the payoff vector \( [Ax] \) is obtained via (2.1):

\[
[Ax] = \begin{pmatrix}
y\delta - z\varepsilon \\
-x\varepsilon + z\delta \\
x\delta - y\varepsilon \\
\end{pmatrix}
\]

Average fitness of the population is:

\[
\mathbf{x}Ax = x(y\delta - z\varepsilon) + y(-x\varepsilon + z\delta) + z(x\delta - y\varepsilon)
\]

The replicator equation (2.3) with the game matrix (2.14) induce on the 2-simplex the following vector field:

\[
\begin{pmatrix}
\dot{x} = x[y\delta - z\varepsilon - (x(y\delta - z\varepsilon) + y(-x\varepsilon + z\delta) + z(x\delta - y\varepsilon))] \\
\dot{y} = y[-x\varepsilon + z\delta - (x(y\delta - z\varepsilon) + y(-x\varepsilon + z\delta) + z(x\delta - y\varepsilon))] \\
\dot{z} = z[x\delta - y\varepsilon - (x(y\delta - z\varepsilon) + y(-x\varepsilon + z\delta) + z(x\delta - y\varepsilon))] \\
\end{pmatrix}
\] (2.15)

While Hofbauer and Sigmund (2003) use the Poincare-Bendixon theorem together with the Dulac criterion to prove that limit cycles cannot occur in games with three strategies under the replicator we will derive this negative result using tools from dynamical systems, in particular the Hopf bifurcation and ‘normal form’ theory.
The same toolkit will be applied next to the Logit Dynamic and a positive result - stable limit cycles do occur - will be derived.

As we are interested in limit cycles within the simplex we consider only interior fixed points of this system (any replicator dynamic has the simplex vertices as steady states, too). For the parameter range $\epsilon, \delta > 0$ the barycentrum $x^* = [x = 1/3, y = 1/3, z = 1/3]$ is always an interior fixed point of (2.15). In order to analyze its stability properties we obtain first, by substituting $z = 1 - x - y$ into (2.15), a proper 2-dimensional dynamical system of the form:

$$
\begin{bmatrix}
\dot{x} \\
\dot{y}
\end{bmatrix} =
\begin{bmatrix}
-x\epsilon + 2xy\epsilon - x^2\delta + 2x^2\epsilon + x^3\delta - x^3\epsilon + xy^2\delta + x^2y\delta - xy^2\epsilon - x^2y\epsilon \\
y\delta - 2xy\delta - 2y^2\delta + y^3\epsilon - y^3\delta + xy^2\delta + x^2y\delta - xy^2\epsilon - x^2y\epsilon
\end{bmatrix}
$$

(2.16)

We can detect the Hopf bifurcation threshold at the point where the trace of the Jacobian matrix of (2.16) is equal to zero. The Jacobian evaluated at the barycentrical steady state $x^*$ is

$$
\begin{bmatrix}
\epsilon & \delta + \epsilon \\
-\delta - \epsilon & -\delta
\end{bmatrix}
$$

with eigenvalues: $\lambda_{1,2}(\epsilon, \delta) = \frac{1}{2}(\epsilon - \delta) \pm i \frac{\sqrt{3}}{2} \sqrt{(\epsilon + \delta)^2}$ and trace $\epsilon - \delta$. For $\delta < \epsilon$, $x^*$ is an unstable focus, at $\delta = \epsilon$ a pair of imaginary eigenvalues crosses the imaginary axis($\lambda_{1,2} = \pm i \sqrt{3}\delta$), while for $\delta > \epsilon$ $x^*$ becomes a stable focus (see Fig. (2.1) below). This is consistent with Theorem 6 in Zeeman (1980) which states that the determinant of the matrix $A$ determines the stability properties of the interior fixed point. In example (2.14) $DetA = \epsilon^3 - \delta^3$ vanishes for $\epsilon = \delta$. By the same Theorem 6 in Zeeman (1980) the vector field (2.15) has a center in the 2-simplex and a continuum of cycles if $DetA = 0$.

Alternatively, our local bifurcation analysis suggests that a Hopf bifurcation occurs when $\epsilon = \delta$ and in order to ascertain its features - sub/supercritical or degenerate - we have to further investigate the nonlinear vector field near the $(x^*, \epsilon = \delta)$ point.

At the Hopf bifurcation $\epsilon = \delta$ necessary condition $Re \lambda_{1,2}(\epsilon, \delta) = 0$, vector field (2.16)
takes a simpler form:
\[
\begin{bmatrix}
\dot{x} \\
\dot{y}
\end{bmatrix} =
\begin{bmatrix}
-x\delta + 2xy\delta + x^2\delta \\
y\delta - 2xy\delta - y^2\delta
\end{bmatrix}
\]

Using formula (2.8) we can back out the nonlinear functions \(f^1, f^2\) needed for the computation of the first Lyapunov coefficient:
\[
\begin{align*}
f_1(x, y) &= y\sqrt{3}\delta - x\delta + 2xy\delta + x^2\delta \\
f_2(x, y) &= -x\sqrt{3}\delta + y\delta - 2xy\delta - y^2\delta
\end{align*}
\]

We can now state the following result:

**Lemma 1** All Hopf bifurcations are degenerate in the circulant Rock-Scissor-Paper game under Replicator Dynamics.

**Proof.** From the nonlinear functions \(f_1(x, y), f_2(x, y)\) derived above we can easily compute the first Lyapunov coefficient (2.10) as \(l_1(\varepsilon^{Hopf} = \delta^{Hopf}) = 0\) which means that there is a first degeneracy in the third order terms from the Taylor expansion of the normal form. The detected bifurcation is a Bautin or degenerate Hopf bifurcation (assuming away other higher order degeneracies: technically, the second Lyapunov coefficient \(l_1(\varepsilon^{Hopf} = \delta^{Hopf})\) should not vanish).

Although, in general, the orbital structure at a degenerate Hopf bifurcation may be extremely complicated (see Section (2.3)), for our particular vector field induced by the Replicator Dynamics it can be shown by Lyapunov function techniques (Hofbauer and Sigmund (2003), Zeeman (1980)) that a continuum of cycles is born exactly at the critical parameter value\(^8\).

\(^8\)The absence of a generic Hopf bifurcation does not suffice to conclude that the vector field admits no isolated periodic orbits and 'global' results are required (e.g. the Bendixson-Dulac method, positive divergence of the vector field on the simplex).
Figure 2.1: Rock-Scissors-Paper and Replicator Dynamics for fixed $\varepsilon = 1$ and varying $\delta$. Qualitative changes in the phase portraits - unstable focus (Panel (a)), continuum of cycles (Panel (b)) and stable focus (Panel (c)) - occur as we increase $\delta$ from below to above $\varepsilon$.

It would be worth pointing out the connection between our local (in)stability results and the static concept of Evolutionary Stable Strategy (ESS). As already noted in the literature (Zeeman (1980), Hofbauer (2000)) ESS implies (global) asymptotic stability under a wide class of evolutionary dynamics - Replicator Dynamics, Best Response and Smooth Best Response Dynamics, Brown-von Newmann-Nash, etc. The reverse implication does not hold in general, i.e. the local stability analysis does not suffice to qualify an attractor as an ESS. However, for $\delta < \varepsilon$ we have shown that $x^*$ is an unstable focus and we can conclude that, for this class of RSP games, the barycentrum is not an ESS. Indeed, the case $\delta < \varepsilon$ is the so called bad RSP game (Sandholm (2008)) which is known not to have an ESS. For $\delta > \varepsilon$ we are in the good RSP game and it does have an interior ESS which coincides with the asymptotically stable rest point $x^*$ of the Replicator Dynamics. Last, if $\delta = \varepsilon$ (i.e. standard RSP) then $x^*$ is a neutrally stable strategy/state and we proved that in this zero-sum game the Replicator undergoes a degenerate Hopf bifurcation. The Hopf bifurcation scenario is illustrated in Fig. 2.1 for a particular value of $\varepsilon$. 

31
Circulant RSP Game and Logit Dynamics

The logit evolutionary dynamics (2.6) applied to our normal form game (2.14) leads to the following vector field:

\[
\begin{align*}
\dot{x} &= \frac{\exp(\beta(y\delta - z\varepsilon))}{\exp(\beta(y\delta - z\varepsilon)) + \exp(\beta(-x\varepsilon - z\delta)) + \exp(\beta(x\delta - y\varepsilon))} - x \\
\dot{y} &= \frac{\exp(\beta(-x\delta + z\varepsilon))}{\exp(\beta(y\delta - z\varepsilon)) + \exp(\beta(-x\varepsilon - z\delta)) + \exp(\beta(x\delta - y\varepsilon))} - y \\
\dot{z} &= \frac{\exp(\beta(x\delta - y\varepsilon))}{\exp(\beta(y\delta - z\varepsilon)) + \exp(\beta(-x\varepsilon - z\delta)) + \exp(\beta(x\delta - y\varepsilon))} - z
\end{align*}
\] (2.17)

By substituting \( z = 1 - x - y \) into (2.17) we can reduce the initial system to a 2-dimensional system of equations and solve for its fixed points in the interior of the simplex, numerically, for different parameters values \((\delta, \varepsilon, \beta)\):

\[
\begin{align*}
\frac{\exp(\beta(y\delta - z(-x-y+1)))}{\exp(\beta(x\delta - y\varepsilon)) + \exp(\beta(-x\delta + z(-x-y+1))) + \exp(\beta(y\delta - (x-y+1)))} - x &= 0 \\
\frac{\exp(\beta(-x\delta + z(-x-y+1)))}{\exp(\beta(x\delta - y\varepsilon)) + \exp(\beta(-x\delta + z(-x-y+1))) + \exp(\beta(y\delta - (x-y+1)))} - y &= 0
\end{align*}
\] (2.18)

The 2 – dim simplex barycentrum \([x = 1/3, y = 1/3, z = 1/3]\) remains a fixed point irrespective of the value of \( \beta \). The Jacobian of (2.18) evaluated at this steady state is:

\[
\begin{bmatrix}
\frac{1}{3} \beta \varepsilon - 1 & \frac{1}{3} \beta \delta + \frac{1}{3} \beta \varepsilon \\
-\frac{1}{3} \beta \delta - \frac{1}{3} \beta \varepsilon & -\frac{1}{3} \beta \delta - 1
\end{bmatrix}
\]

with eigenvalues: \( \lambda_{1,2} = \frac{1}{6} \beta \varepsilon - \frac{1}{6} \beta \delta - 1 \pm i \frac{1}{3} \sqrt{1 - \frac{1}{3}(\beta \delta + \beta \varepsilon)} \).

The Hopf bifurcation (necessary) condition \( \text{Re}(\lambda_{1,2}) = 0 \) leads to:

\[
\beta^{\text{Hopf}} = 6 \varepsilon - \delta, \quad 0 < \delta < \varepsilon
\] (2.19)

We notice that for the zero-sum RSP game \((\varepsilon = \delta)\) - unlike Replicator Dynamics which exhibited a degenerate Hopf at \( \varepsilon = \delta \) - the barycentrum is always asymptotically stable (\( \text{Re} \lambda_{1,2} = -1 \)) under Logit Dynamics. We have the following:

**Lemma 2** The Logit Dynamics (2.17) on the circulant Rock-Scissors-Paper game exhibits a generic Hopf bifurcation and, therefore, has limit cycles. Moreover, all such Hopf bifurcations are supercritical, i.e. the limit cycles are born stable.
Proof. Condition (2.19) gives the necessary first-order condition for Hopf bifurcation to occur; in order to show that the Hopf bifurcation is non-degenerate we have to compute, according to (2.10) the first Lyapunov coefficient \( l_1(\beta^{Hopf}, \varepsilon, \delta) \) and check whether it is non-zero. The analytical form of this coefficient takes a complicated expression of exponential terms (see appendix (2.A)) which, after some tedious computations, boils down to:

\[
l_1(\beta^{Hopf}, \varepsilon, \delta) = -\frac{864(\delta \varepsilon + \delta^2 + \varepsilon^2)}{3(3 \varepsilon - 3 \delta)^2} < 0, \quad \varepsilon > \delta > 0.
\]

Computer simulations of this route to a stable cycle are shown in Fig. 2.2 below. We notice that as \( \beta \) moves up from 10 to 35 (i.e. the noise level is decreasing) the interior stable steady state loses stability via a supercritical Hopf bifurcation and a small, stable limit cycle emerges around the unstable steady state. Unlike Replicator Dynamics, stable cyclic behavior does occur under the Logit dynamics even for three-strategy games.

![Figure 2.2](image_url)

(a) Stable focus, \( \beta = 10 \)  \hspace{1cm} (b) Generic Hopf, \( \beta = 30 \)  \hspace{1cm} (c) Limit cycle, \( \beta = 35 \)

Figure 2.2: Rock-Scissors-Paper and Logit Dynamics for fixed game \( \varepsilon = 1, \delta = 0.8 \) and free behavioral parameter \( \beta \). Qualitative changes in the phase portraits: a stable interior fixed point (Panel (a)) loses stability when the critical threshold \( \beta = 30 \) is hit, via a generic, supercritical Hopf bifurcation (Panel (b)) and, if \( \beta \) is pushed up even further, a stable limit cycle is born (Panel (c)).
Similar periodic behaviour can be detected in the payoff parameter space as in Fig. 2.3 below where the noise level is kept constant and the game is allowed to change.

Figure 2.3: Rock-Scissors-Paper and Logit Dynamics for fixed behavioral parameter $\beta = 10$ and free game parameter $\delta [\varepsilon = 1]$. Qualitative changes in the phase portraits: a stable interior fixed point (Panel (a)) loses stability when the critical threshold $\delta = 0.399$ is hit, via a generic, supercritical Hopf bifurcation (Panel (b)). Panel (c) display a stable limit cycle for $\delta = 0.1$ that attracts trajectories originating both outside and inside the cycle.

A more thorough search in the parameters space is performed with the help of a sophisticated continuation software\(^9\) (Dhooge et al. (2003)) which enable us to determine all combination of relevant parameters giving rise to a specific singularity. Figure 2.4 depicts curves of Hopf bifurcations in the $(\beta, \varepsilon)$ and $(\beta, \delta)$ parameter space (fixing $\delta(\varepsilon)$ to $1$). As we cross these Hopf curves from below the stable interior fixed point loses stability and a stable periodic attractor surrounds it.

\(^9\)MatCont (see package documentation at http://sourceforge.net/projects/matcont/).
Generalized RSP Game and Logit Dynamics

In this subsection we discuss the occurrence of limit cycles in the most general specification of a Rock-Scissors-Paper game (2.13) under Logit Dynamics. For circulant RSP game (i.e. circulant payoff matrix) Weissing (1991) proved that the discrete and continuous time Replicator Dynamics are 'qualitatively' equivalent. However, for a non-circulant payoff matrix (2.20), co-existing stable and unstable limit cycles are found under discrete-time Replicator Dynamics. This sharply contrasts with the behaviour of the continuous-time Replicator Dynamics which is known not to give rise to generic Hopf bifurcations. On the other hand, as shown in the previous subsection stable limit cycles do occur under Logit Dynamics so a natural question to pose is whether multiple, interior periodic attractors can be detected under the logistic dynamics in the generalized version of the circulant RSP game.

Weissing (1991) presents the following generalized RSP game as an example of co-existing limit cycles under discrete-time Replicator Dynamics:
\[ W = \begin{pmatrix} k & b_1 & d_1 \\ d_2 & k & b_2 \\ b_1 + 3\mu & d_1 - 3\mu & k \end{pmatrix}, \quad b_i > k \geq d_i, d_1 \geq 3\mu > 0 \quad (2.20) \]

We have already seen that this general specification can be turned into (2.13) by a positive linear payoff transformations as the continuous-time dynamics remain invariant under such transformation\(^{10}\).

In the sequel we will investigate, numerically, the dynamical system resulting from the generalized game payoff matrix (2.13) and the Logit evolutionary dynamics:

\[ \dot{x}_i = \frac{e^{\beta(Ax)_i}}{\sum_{i=1}^{3} e^{\beta(Ax)_i}}, \quad A x = \begin{bmatrix} \delta_2x_2 - \varepsilon_3x_3 \\ \delta_3x_3 - \varepsilon_1x_1 \\ \delta_1x_1 - \varepsilon_2x_2 \end{bmatrix} \]

Figure 2.5 displays typical curves of fold and Hopf bifurcations produced by MatCont, along with their codimension II singularities in various sub-regions of the parameters space. The curves are initiated at certain bifurcation points\(^{11}\) detected along the system’s equilibria curves. They first generalize results obtained in previous subsection for a circulant RSP payoff matrix to the entire class of RSP games. Second, from the bifurcation diagrams we can detect, the co-existence of multiple interior steady-states (born via Limit Point bifurcation) and limit cycle (created when a Hopf curve is crossed). However, the fold curves seem to emerge beyond the generalized RSP class of games (i.e. the positive \(\mathbb{R}^2\) orthant in each 2-parameter subspace)

In order to address the question of co-existing limit cycles under continuous-time

---

\(^{10}\)For continuous-time Replicator Dynamics the proof could be found in Weissing (1991)). Yet, the dynamics on (2.20 ) and (2.13) may be qualitatively different under discrete-time version of the Replicator Dynamics.

\(^{11}\)A supercritical (first Lyapunov coefficient \(l_1 = -3.222\)) Hopf point is detected at \(\delta_1 = 0.626\) and two Fold points at \(\delta_1 = -0.554\) and \(\delta_1 = -0.952\). Benchmark parameterization: \(\beta = 8.3, \delta_1 = 4, \delta_2 = 3, \delta_3 = 2, \varepsilon_1 = 2, \varepsilon_21 = 3, \varepsilon_3 = 4\).
Logit Dynamics, one should run a systematic search in the parameter space for a fold bifurcation of limit cycles: detect the Hopf points, continue them with respect to each of the 6-parameter in the game form (2.13) and, then track any codimension II singularity (most interestingly, the limit point of cycles bifurcation). A representative subset of the resulting Hopf curves is plotted in Fig. 2.5. Except for the Bogdanov-Takens (BT) bifurcation, no further codimension II singularities are detected by the continuation package in the positive $\mathbb{R}^2$ orthant corresponding to the generalized class of RSP game. In particular, there is no limit point/fold bifurcation of cycles, i.e. multiple, co-existing (un)stable limit cycles are ruled out.
Figure 2.5: Generalized Rock-Scissors-Paper - curves of codimension I - fold (LP) and Hopf (H) - bifurcations along with the detected codimension II singularities - Cusp (CP), Bogdanov-Taken (BT), Cusp (CP) and Zero-Hopf (ZH) points - in various two-parameter subspaces. Unless free to float, game parameters are fixed to following values - $\delta_1 = 0.62$, $\delta_2 = 4$, $\delta_3 = 3$, $\varepsilon_1 = 2$, $\varepsilon_2 = 3$, $\varepsilon_3 = 4$ - while the intensity of choice is $\beta = 8$. 


### 2.4.2 Coordination Game

Using topological arguments, Zeeman (1980) shows that three-strategies games have at most one interior, isolated fixed point under Replicator Dynamics (see Theorem 3 in Zeeman (1980)). In particular, fold catastrophe (two isolated fixed points which collide and disappear when some parameter is varied) cannot occur within the simplex. In this section we provide - by means of the classical coordination game - numerical evidence for the occurrence of multiple, isolated interior steady-states under the Logit Dynamics and show that fold catastrophe is possible when we alter the intensity of choice $\beta$. The coordination game we consider is given by the following payoff matrix:

$$A = \begin{pmatrix}
1 - \varepsilon & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 + \varepsilon
\end{pmatrix}, \varepsilon \in (0,1)$$

#### Coordination Game and Replicator Dynamics

The vector field $\dot{x} = x[Ax - xA]$, becomes:

$$\begin{cases}
\dot{x} = x[x (-\varepsilon + 1) - (y^2 + z^2 (\varepsilon + 1) + x^2 (-\varepsilon + 1))] \\
\dot{y} = y[y - (y^2 + z^2 (\varepsilon + 1) + x^2 (-\varepsilon + 1))] \\
\dot{z} = z[z (\varepsilon + 1) - (y^2 + z^2 (\varepsilon + 1) + x^2 (-\varepsilon + 1))]
\end{cases} \quad (2.21)$$

This systems has the following fixed points in simplex coordinates:

(i) simplex vertices (stable nodes): $(1,0,0), (0,1,0), (0,0,1)$

(ii) interior (source): $O \left( \frac{1 + \varepsilon}{3 - \varepsilon^2}, \frac{1 - \varepsilon^2}{3 - \varepsilon^2}, \frac{1 - \varepsilon}{3 - \varepsilon^2} \right)$

(iii) on the boundary (saddles):

$M \left( \frac{1}{2 - \varepsilon}, \frac{1 - \varepsilon}{2 - \varepsilon}, 0 \right), N \left( \frac{1}{2 - \varepsilon}, 0, \frac{1 - \varepsilon}{2 - \varepsilon} \right), P \left( 0, \frac{1}{2 - \varepsilon}, \frac{1 - \varepsilon}{2 - \varepsilon} \right)$

The eigenvalues of the corresponding Jacobian evaluated at each of the above fixed points are:

$A [x = 0, y = 1, z = 0]$; eigenvalues: $\lambda_{1,2,3} = -1$
$B \ [x = 1, y = 0, z = 0]$; eigenvalues: $\lambda_{1,2,3} = \varepsilon - 1$

$C \ [x = 0, y = 0, z = 1]$; eigenvalues: $\lambda_{1,2,3} = -\varepsilon - 1$

$O \ [x = \frac{\varepsilon + 1}{3-\varepsilon^2}, y = \frac{1 - \varepsilon^2}{3-\varepsilon^2}, z = \frac{1 - \varepsilon}{3-\varepsilon^2}]$; eigenvalues: $\lambda_{1,2} = \frac{\varepsilon^2 - 1}{\varepsilon^2 - 3} > 0$, for $\varepsilon \in (0, 1)$

$M \ [x = \frac{1}{2 - \varepsilon}, y = \frac{1 - \varepsilon}{2 - \varepsilon}, z = 0]$; eigenvalues: $\lambda_{1,2} = \pm \frac{1 - \varepsilon}{2 - \varepsilon}$

$N \ [x = \frac{1}{2 - \varepsilon}, y = 0, z = \frac{1 - \varepsilon}{2 - \varepsilon}]$; eigenvalues: $\lambda_{1,2} = \pm \frac{1}{2 - \varepsilon} (2\varepsilon + \varepsilon^2 - \varepsilon^3 - 2)$

$P \ [x = 0, y = \frac{1}{2 - \varepsilon}, z = \frac{1 - \varepsilon}{2 - \varepsilon}]$; eigenvalues: $\lambda_{1,2} = \pm \frac{1}{4\varepsilon + \varepsilon^2 + 4} (\varepsilon + \varepsilon^2 - \varepsilon^3 - 2)$

This fixed points structure is consistent with Zeeman (1980) result that no fold catastrophes (i.e. multiple, isolated and interior steady states) can occur under the Replicator Dynamics. The three saddles together with the interior source define the basins of attractions for the simplex vertices. The boundaries of the simplex are invariant under Replicator Dynamics\textsuperscript{12} and it suffices to show that the segment lines $[OM], [ON]$, and $[OP]$ (see Fig. 2.6 for a plot of these segments for $\varepsilon = 0.1$) are also invariant under this dynamics.

Figure 2.6: Coordination Game with $\varepsilon = 0.1$ and Replicator Dynamics - Invariant manifolds

\textsuperscript{12}see proof in Hofbauer and Sigmund (2003) pp. 67-68.
Lemma 3  The segment lines \([OM], [ON],\) and \([OP]\) are invariant under Replicator Dynamics and they form, along with the simplex edges, the boundaries of the basins of attraction for the three stable steady states.

Proof. Using the standard substitution \(z = 1 - x - y\) system (2.21) becomes:

\[
\begin{bmatrix}
\dot{x} \\
\dot{y}
\end{bmatrix}
= \begin{bmatrix}
x^2 - x - xy - x^2\varepsilon - x(-x - y + 1)^2 - x\varepsilon(-x - y + 1)^2 \\
y^2 - y^3 - x^2y + x^2y\varepsilon - y(-x - y + 1)^2 - y\varepsilon(-x - y + 1)^2
\end{bmatrix}
\]

Segment \([MO]\) is defined, in simplex coordinates, by \(y = x(1 - \varepsilon)\). Along this line we have:

\[
\begin{bmatrix}
\dot{y} \\
\dot{x}
\end{bmatrix}_{[MO]} = \frac{x(-\varepsilon + 1)(x(-\varepsilon + 1) - x^2(-\varepsilon + 1) - x^2(-\varepsilon + 1)^2 - (\varepsilon + 1)(-x - y + 1)^2)}{x(x(-\varepsilon + 1) - x^2(-\varepsilon + 1) - x^2(-\varepsilon + 1)^2 - (\varepsilon + 1)(-x - y + 1)^2)}
= 1 - \varepsilon,
\]

which is exactly the slope of \([MO]\). Similarly, invariance results are obtained for segments \([ON]\) and \([OP]\) defined by \(x(1 - \varepsilon) = z(1 + \varepsilon)\) and \(y = z(1 + \varepsilon)\). ■

Analytically, the basins of attraction are determined by the areas of the polygons delineated by the invariant manifolds \([OM], [ON],\) and \([OP]\) and the 2-dim simplex boundaries, as follows:

\[
\mathcal{A}(1, 0, 0) = \mathcal{A}[AMON] = \begin{bmatrix}
\frac{1}{8} \sqrt{3} \frac{\varepsilon + 1}{3(2 - \varepsilon)(3 - \varepsilon)} - \frac{1}{4} \sqrt{\frac{3}{2}} \frac{3 - \varepsilon}{6 - 2 \varepsilon} + \frac{1}{4} \sqrt{\frac{3}{2}} \frac{3 - \varepsilon}{6 - 2 \varepsilon} - \frac{1}{4} \sqrt{\frac{3}{2}} \frac{3 - \varepsilon}{6 - 2 \varepsilon} \\
+ \frac{1}{4} \sqrt{\frac{3}{2}} \frac{\varepsilon + 1}{3(2 - \varepsilon)(3 - \varepsilon)} + \frac{1}{4} \sqrt{\frac{3}{2}} \frac{3 - \varepsilon}{6 - 2 \varepsilon} - \frac{1}{4} \sqrt{\frac{3}{2}} \frac{3 - \varepsilon}{6 - 2 \varepsilon} \\
+ \frac{1}{4} \sqrt{\frac{3}{2}} \frac{3 - \varepsilon}{6 - 2 \varepsilon} + \frac{1}{4} \sqrt{\frac{3}{2}} \frac{\varepsilon + 1}{3(2 - \varepsilon)(3 - \varepsilon)} - \frac{1}{4} \sqrt{\frac{3}{2}} \frac{3 - \varepsilon}{6 - 2 \varepsilon}
\end{bmatrix}
\]

\[
\mathcal{B}(0, 1, 0) = \mathcal{A}[BMOP] = \begin{bmatrix}
\frac{1}{4} \sqrt{3} \frac{\varepsilon + 1}{3(2 - \varepsilon)(3 - \varepsilon)} - \frac{1}{4} \sqrt{\frac{3}{2}} \frac{3 - \varepsilon}{6 - 2 \varepsilon} \\
+ \frac{1}{4} \sqrt{\frac{3}{2}} \frac{\varepsilon + 1}{3(2 - \varepsilon)(3 - \varepsilon)} - \frac{1}{4} \sqrt{\frac{3}{2}} \frac{3 - \varepsilon}{6 - 2 \varepsilon}
\end{bmatrix}
\]

\[
\mathcal{C}(0, 0, 1) = \mathcal{A}[CNOP] = \begin{bmatrix}
\frac{1}{4} \sqrt{3} \frac{(\varepsilon + 1)(1 - \varepsilon) - (1 + \varepsilon)}{2(2 - \varepsilon)(3 - \varepsilon)} \\
+ \frac{1}{4} \sqrt{\frac{3}{2}} \frac{\varepsilon + 1}{3(2 - \varepsilon)(3 - \varepsilon)} - \frac{1}{4} \sqrt{\frac{3}{2}} \frac{3 - \varepsilon}{6 - 2 \varepsilon}
\end{bmatrix}
\]

The basins’ sizes vary with the payoff perturbation parameter \(\varepsilon\) as in Table 2.1.

When the payoff asymmetries are small, the simplex of initial conditions is divided equally among the three fixed points/equilibria (Fig. 2.7a). As the payoff discrepancies increase most of the initial conditions are attracted by the welfare-maximizing equilibrium \((0, 0, 1)\) (see Fig. 2.7b).
<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$A(1,0,0)$</th>
<th>$B(0,1,0)$</th>
<th>$C(0,0,1)$</th>
<th>Long-Run Average Fitness/Welfare</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>33.33%</td>
<td>33.33%</td>
<td>33.33%</td>
<td>1</td>
</tr>
<tr>
<td>0.1</td>
<td>29.9%</td>
<td>33.2%</td>
<td>36.7%</td>
<td>1.0048</td>
</tr>
<tr>
<td>0.5</td>
<td>15.5%</td>
<td>30.3%</td>
<td>54.54%</td>
<td>1.1986</td>
</tr>
<tr>
<td>0.6</td>
<td>8.92%</td>
<td>12.18%</td>
<td>78.90%</td>
<td>1.4199</td>
</tr>
<tr>
<td>0.9</td>
<td>1.2%</td>
<td>12%</td>
<td>86.7%</td>
<td>1.512</td>
</tr>
</tbody>
</table>

Table 2.1: Coordination Game, Replicator Dynamics–Long-run average fitness for different payoff-perturbation parameter $\varepsilon$

![Figure 2.7](image)

Figure 2.7: Coordination Game and Replicator Dynamics-basins of attraction for different values of the payoff parameter $\varepsilon$

**Coordination Game and Logit Dynamics**

We choose a small payoff perturbation, $\varepsilon = 0.1$, such that system is ‘close’ to the symmetric basins of attraction $A, B$ and $C$ in the Replicator Dynamics. Logit Dynamics together with payoff matrix $A$ define the following vector field on the simplex of frequencies $(x, y, z)$ of strategies $E_1, E_2, E_3$, respectively:
\[
\begin{align*}
\dot{x} &= \frac{\exp(0.9x)}{\exp(0.9x) + \exp(\beta y) + \exp(1.1z)} - x \\
\dot{y} &= \frac{\exp(\beta y)}{\exp(0.9x) + \exp(\beta y) + \exp(1.1z)} - y \\
\dot{z} &= \frac{\exp(1.1z)}{\exp(0.9x) + \exp(\beta y) + \exp(1.1z)} - z
\end{align*}
\]

In order to ascertain the number of (asymptotically) stable fixed points on the two-dimensional simplex we first run simulations for increasing values of $\beta$ and for different initial conditions within the simplex:

Case I. $\beta \approx 0$. One interior steady state (initial conditions, time series and attracting point in Fig. 2.8a). The barycentrum is globally attracting, i.e. irrespective of the initial proportions, the population will settle down to a state with equalized fractions.

Case II. $\beta \geq 10$. Three stable steady states asymptotically approaching the vertices of the simplex as $\beta$ increases (Fig. 2.8b-d) plus other unstable steady states. The size of their basins of attraction is determined both by the strategy relative payoff advantage and by the value of the intensity of choice. Whichever of the three strategy happens to outnumber the other two in the initial population, will eventually win the evolutionary competition.
(a) \( \beta = 0 \), any initial mixture \((x, y, z)\) converges to \((1/3, 1/3, 1/3)\)

(b) \( \beta = 10 \), initial mixture \((0.4, 0.3, 0.3) \to (1, 0, 0)\)

(c) \( \beta = 10 \), initial mixture \((0.25, 0.4, 0.35) \to (0, 1, 0)\)

(d) \( \beta = 10 \), initial fractions \((0.33, 0.33, 0.34) \to (0, 0, 1)\)

Figure 2.8: Coordination Game\(\varepsilon = 0.1\) and Logit Dynamics - Unique stable steady state for \( \beta \) low (Panel(a)) and three, co-existing stable steady states for high \( \beta = 10 \) (Panels (b)-(d)).

**Multiple Equilibria.** The most interesting situations occur in transition from high to low values of the behavioral parameter \( \beta \): we will show that the Logit Dynamics undertakes a sequence of fold bifurcations by which each of the three stable
steady states collides with another unstable steady state from within the simplex and disappears. Fig. 2.9cd shows simulations of time series converging to a unique steady state for low values of $\beta$ ($\beta = 2.6$). Panels (a)-(b) show the co-existence of two steady-states for $\beta = 3$ as different initial population mixtures converge to different steady states.

Figure 2.9: Coordination Game[$\varepsilon = 0.1$] and Logit Dynamics with moderate level of noise
-Panels (a)-(b) display two interior, isolated stable steady states, while Panels (c)-(d) shows convergence to a unique, interior steady state
**Bifurcations.** Unlike Replicator, the Logit Dynamics displays multiple, interior isolated steady states created via a fold bifurcation. In the particular case of a 3-strategy Coordination game, three interior stable steady states emerge through a sequence of two saddle-node bifurcations.

Fig. 2.10 depicts the fold bifurcations scenario by which the multiple, interior fixed points appear when the intensity of choice (Panel (a)) or the payoff parameter (Panel (b)) changes. For small values of $\beta$ the unique, interior stable steady state is the simplex barycentrum $(1/3, 1/3, 1/3)$. As $\beta$ increases this steady state moves towards the Pareto-superior equilibrium $(0, 0, 1)$. A first fold bifurcation occurs at $\beta = 2.77$ and two new fixed points are created, one stable and one unstable. If we increase $\beta$ even further ($\beta \approx 3.26$) a second fold bifurcation takes place and two additional equilibria emerge, one stable and one unstable. Last, two new fixed points arise at $\beta = 4.31$ via a saddle-source bifurcation$^{13}$. There is co-existence of three stable steady states for large values of the intensity of choice $\beta$, the ‘logit equilibria’. A similar sequence of bifurcations is visible in the payoff parameter space (Fig. 2.10b) where a family of fold bifurcations is obtained for a particular value of the switching intensity $\beta$.

$^{13}$At a saddle-source bifurcation two additional unstable steady states emerge.
Figure 2.10: Coordination Game and Logit Dynamics. Curves of equilibria along with codimension I singularities, in this case fold (LP) points. Panel (a)-(b): the multiplicity of steady states arises through a sequence of three limit point (fold) bifurcations as one of the two parameters varies.

The continuation of the fold curves in the \((\beta, \varepsilon)\) parameter space allows the detection of codimension II bifurcations. The cusp points in Fig. 2.11 below organize the entire bifurcating scenario and capture the emergence of multiple, interior point-attractors in Coordination Game under the smoothed best response dynamics. Once such a curve of fold points is crossed from below two additional equilibria are created: one stable and one unstable after a saddle-node bifurcation and two unstable steady states after a saddle-source bifurcation. If choice is virtually random \((\beta = 0)\) there is an unique steady state while for large \(\beta\) there are 7 steady states (3 stable and 4 unstable).
Basins of Attraction. The numerical computation of the basins of attraction for different equilibria reveals interesting properties of the Logit dynamics from a social welfare perspective. We construct a measure of long-run aggregate welfare as the payoff at the stable steady state weighted by its corresponding basin of attraction size. While for extreme values of the intensity of choice the basins of attraction are similar in size with the Replicator Dynamics (Panels (a), (c), (d) in Fig. 2.12,2.13), for moderate levels of rationality the population manages to coordinate close to the Pareto optimal Nash equilibria (Panel (b) in Fig. 2.12,2.13). The aggregate welfare evolves non-monotonically with respect to the behavioural parameter $\beta$ and, for a given payoff perturbation $\varepsilon$, it is maximal just before the first bifurcation occurs $\beta_{LP_1}^{LF} = 2.77$. 

Figure 2.11: Coordination Game and Logit Dynamics. Curves of fold points along with detected codimension II singularities - in this case, cusp (CP) and zero-Hopf (ZH) points - traced in the $(\varepsilon, \beta)$ parameter space
(a) $\beta = 1$, one barrycentrical steady state

(b) $\beta = 10/4$, one steady state ‘close’ to the Pareto superior equilibrium

(c) $\beta = 10/3$, two stable steady states

(d) $\beta = 15$, three stable steady states

Figure 2.12: Coordination Game $[\varepsilon = 0.1]$ and Logit Dynamics: Panels (a)-(d), basins of attraction for increasing values of the intensity of choice $\beta$. Fractions converging to each of the steady state are indicated in the box.
Figure 2.13: Coordination Game $[\varepsilon = 0.6]$ and Logit Dynamics: Panels (a)-(d), basins of attraction for increasing values of the intensity of choice $\beta$

The long run average welfare evolves non-monotonically with respect to the behavioural parameter $\beta$ (Fig. 2.14ab): it increases as the fully mixed equilibrium slides slowly towards the Pareto optimal $(0, 0, 1)$ vertex, attains a maximum before the first fold bifurcation occurs at $\beta^{LP_1} = 2.77$ and then decreases, approaching the Replicator Dynamics average welfare, in the limit of $\beta \to \infty$. There are two effects driving the welfare peak before $\beta^{LP_1}$ is hit: first, the steady state payoff is higher the
closer the steady state is to the Pareto optimal equilibrium. Second there is a ‘basin of attraction’ effect: for $\beta < \beta^{LP}$ the entire simplex of initial conditions is attracted by the unique steady state lying close to the optimal equilibrium. Intuitively, the noisy choices in the low-beta regime help players escape the path-dependence built into Coordination games while for super-rational play are trapped into best, albeit payoff-inferior, responses.

In the limiting case $\beta \to \infty$, the fixed points of the Logit Dynamics (i.e. the logit equilibria) coincide with the Nash equilibria of the underlying game which, for this Coordination Game, are exactly the fixed points of the Replicator Dynamics. Thus the analysis (stable fixed points, basins’ of attraction sizes) of the ‘unbounded’ rationality case is identical to the one pertaining to the Replicator Dynamics in Coordination game (see subsection (2.4.2)).

![Figure 2.14](image)

(a) $W - \beta$ plot for different $\varepsilon$

(b) surface plot $(W, \beta, \varepsilon)$

Figure 2.14: Coordination Game and Logit Dynamics-Long-run average welfare plots as function of payoff perturbation parameter $\varepsilon$
2.5 Weighted Logit Dynamics (wLogit)

In this last section we run computer simulations for a three-strategy - Rock-Scissor-Paper - and a four-strategy - Schuster et al. (1991) example-game, from the perspective of a different type of evolutionary dynamics closely related to the Logit dynamic, namely the frequency-weighted Logit:

\[
\dot{x}_i = \frac{x_i \exp[\beta A x]_i}{\sum_k x_k \exp[\beta A x]_k} - x_i, \quad \beta = \eta^{-1}
\]  (2.23)

This evolutionary dynamic has the appealing property that when \( \beta \) approaches 0 it converges to the Replicator Dynamics (with adjustment speed scaled down by a factor \( \beta \)) and when the intensity of choice is very large it approaches the Best Response dynamic (Hofbauer and Weibull (1996)).

2.5.1 Rock-Scissors-Paper and wLogit Dynamics

Dynamic (2.23) together with game payoff matrix (2.14) give rise to the following dynamical system:

\[
\begin{align*}
\dot{x} &= \frac{x \exp[2 \beta e - y \delta]}{x \exp[2 \beta e - y \delta] + y \exp[2 \beta e - z \delta] + z \exp[2 \beta e - x \delta]} - x \\
\dot{y} &= \frac{y \exp[2 \beta e - z \delta]}{x \exp[2 \beta e - y \delta] + y \exp[2 \beta e - z \delta] + z \exp[2 \beta e - x \delta]} - y \\
\dot{z} &= \frac{z \exp[2 \beta e - x \delta]}{x \exp[2 \beta e - y \delta] + y \exp[2 \beta e - z \delta] + z \exp[2 \beta e - x \delta]} - z
\end{align*}
\]

To contrast Replicator and the Logit dynamics with the help of the wLogit we can start by fixing two parameters \( \beta = 5 \) and \( \varepsilon = 1 \), and let \( \delta \) free. We obtain that the barycentrum is an interior fixed point of the system, irrespective of the value of \( \delta \). When the critical threshold \( \delta_{\text{Hopf}} = 1 \) is hit a supercritical Hopf bifurcation arises (first Lyapunov coefficient = -1.249) and a stable limit cycle is born exactly as in the Logit dynamic (cf. Fig. 2.15ab). However, unlike Logit, fluctuations are not limited in amplitude by some Shapley polygon, but they can easily approach the boundary
of the simplex (Fig. 2.15cd) a feature reminiscent of Replicator Dynamics\textsuperscript{14}. The
limit cycles are stable because they are produced by a generic Hopf bifurcation. With
Replicator Dynamics the Hopf bifurcations are degenerate and there is a continuum
of periodic orbits. These features are illustrated in Fig. 2.15 where stable cycles
are detected and continued in the game and behavioral parameters. For, instance,
one family of limit cycles parameterized by the payoff matrix parameter $\varepsilon$ that gets
closer to the simplex boundaries as $\varepsilon$ decreases is shown in Panel (c).

\textsuperscript{14}Notice the difference with the Replicator Dynamics where system converges to a \textit{heteroclinic}
cycle from each simplex vertex to the other.
(a) $\beta = 5, \varepsilon = 1$. Line of equilibria, free $\delta; \delta^{Hopf} = 1$.

(b) $\beta = 5, \varepsilon = 1$. Phase portrait, $\delta^{Hopf} = 1$.

(c) Curve of limit cycles $(x, y, \varepsilon)$.
$\beta = 5, \delta = \delta^{Hopf}$

(d) Curve of limit cycles $(x, y, \beta)$.
$\varepsilon = 1, \delta = \delta^{Hopf}$

Figure 2.15: RSP Game and Weighted Logit Dynamics. Panels (a)-(d): long-run behavior inherits traits from both Logit (stable limit cycles, Panels(a)-(b)) and Replicator Dynamics (cycles of large amplitude, nearing the simplex boundaries, Panels (c)-(d)). Unless free, parameters set to $\varepsilon = 1, \beta = 5, \delta = 1$.
2.5.2 Schuster et al.(1991) Game and wLogit Dynamics

Schuster et al. (1991) derive the following payoff matrix from models of biological interaction:

\[
A = \begin{pmatrix}
0 & 0.5 & -0.1 & 0.1 \\
1.1 & 0 & -0.6 & 0 \\
-0.5 & 1 & 0 & 0 \\
1.7 + \mu & -1 - \mu & -0.2 & 0 \\
\end{pmatrix}
\]  

The Weighted Logit (2.23) with the underlying game (2.24) generates the following vector field on the 4-simplex:

\[
\begin{align*}
\dot{x} &= \frac{x e^{[\beta(0.1t+0.5y-0.1z)]}}{x e^{[\beta(0.1t+0.5y-0.1z)]} + ye^{[\beta(1.1z-0.6x)]} + ze^{[\beta(-0.5x+y)]} + te^{[\beta(-0.2z+y(-\mu-1)+x(\mu+1.7))]}}, \\
\dot{y} &= \frac{ye^{[\beta(1.1z-0.6x)]}}{x e^{[\beta(0.1t+0.5y-0.1z)]} + ye^{[\beta(1.1z-0.6x)]} + ze^{[\beta(-0.5x+y)]} + te^{[\beta(-0.2z+y(-\mu-1)+x(\mu+1.7))]}}, \\
\dot{z} &= \frac{ze^{[\beta(-0.5x+y)]}}{x e^{[\beta(0.1t+0.5y-0.1z)]} + ye^{[\beta(1.1z-0.6x)]} + ze^{[\beta(-0.5x+y)]} + te^{[\beta(-0.2z+y(-\mu-1)+x(\mu+1.7))]}}, \\
\dot{t} &= \frac{te^{[\beta(-0.2z+y(-\mu-1)+x(\mu+1.7))]} }{x e^{[\beta(0.1t+0.5y-0.1z)]} + ye^{[\beta(1.1z-0.6x)]} + ze^{[\beta(-0.5x+y)]} + te^{[\beta(-0.2z+y(-\mu-1)+x(\mu+1.7))]}}. \\
\end{align*}
\]

Schuster et al. (1991) show that, within a specific payoff parameter region \([-0.2 < \mu < -0.105]\), a Feigenbaum sequence of period doubling bifurcations unfolds under the Replicator Dynamics and, eventually, chaos sets in. Here we consider the question whether the weighted version of the Logit Dynamics displays similar patterns, for a given payoff perturbation \(\mu\), when the intensity of choice \(\beta\) is varied. First we fix \(\mu\) to \(-0.2\) (the value for which the Replicator generates 'only' periodic behaviour in Schuster et al. (1991)) and run computer simulations for different \(\beta\). Cycles of increasing period multiplicity are reported in Fig. 2.16.
Figure 2.16: Schuster et al. (1991) game and iLogit Dynamics. Period-Doubling route to chaos. Panels (a)-(d): cycles of increasing period are obtained if the intensity of choice $\beta$ increases.

For $\beta = 2.5$ the wLogit Dynamics already enters the chaotic regime on a strange attractor (Fig. 2.17).
Figure 2.17: Schuster et. al. (1991) game and iLogit Dynamics-Strange Attractor. Panels (a)-(d) projections of the strange attractor onto various two-dimensional state subspaces. Parameters set to $\mu = -0.2, \beta = 2.5$. 
2.6 Conclusions

The main goal of this Chapter was to show that, even for ‘simple’ three-strategy games, periodic attractors do occur under a rationalistic way of modelling evolution in games, the Logit dynamics. The resulting dynamical systems were investigated with respect to changes in both the payoff and behavioral parameters. Identifying stable cyclic behaviour in such a system translates into proving that a generic, non-degenerate Hopf bifurcation occurs. By means of normal form computations, we showed first that a non-degenerate Hopf can not occur for Replicator Dynamics, when the number of strategies is three for games like Rock-Scissors-Paper. In these games, under the Replicator Dynamics, only a degenerate Hopf bifurcation can occur. However, in Logit dynamics, even for three strategy case, stable cycles are created, via a generic, non-degenerate, supercritical Hopf bifurcation. Another finding is that the periodic attractors can be generated either by varying the payoff parameters $(\varepsilon, \delta)$ or the intensity of choice $(\beta)$. Moreover, via computer simulations on a Coordination game, we showed that the Logit may display multiple, isolated, interior steady states together with the fold catastrophe, a bifurcation which is also known not to occur under the Replicator. Interestingly, a measure of aggregate welfare reaches a maximum only for intermediate values of $\beta$, when most of the population manages to coordinate close to the Pareto-superior equilibrium. Last, in a frequency-weighted version of Logit dynamics and for a $4 \times 4$ game, period-doubling route to chaos along with strange attractors emerged when the intensity of choice took moderate, ‘boundedly rational’ values.
2.A Rock-Scissor-Paper game with Logit Dynamics: Computation of the first Lyapunov coefficient

In order to discriminate between a degenerate and a non-degenerate bifurcations we need to compute the first Lyapunov coefficient. For this, we first use equations (2.8) to obtain the nonlinear functions:

\[
\begin{align*}
  f_1(x, y) &= y\sqrt{3}\frac{\varepsilon + \delta}{\varepsilon - \delta} - x + \frac{\exp\left(\frac{6(y\delta - \varepsilon(-x-y+1))}{-\delta + \varepsilon}\right)}{\exp\left(\frac{6(x\delta - y\varepsilon)}{-\delta + \varepsilon}\right) + \exp\left(\frac{6(-x\varepsilon + \delta(-x-y+1))}{-\delta + \varepsilon}\right) + \exp\left(\frac{6(y\delta - \varepsilon(-x-y+1))}{-\delta + \varepsilon}\right)} \\
  f_2(x, y) &= -x\sqrt{3}\frac{\varepsilon + \delta}{\varepsilon - \delta} - y + \frac{\exp\left(\frac{6(x\delta - y\varepsilon)}{-\delta + \varepsilon}\right)}{\exp\left(\frac{6(x\delta - y\varepsilon)}{-\delta + \varepsilon}\right) + \exp\left(\frac{6(-x\varepsilon + \delta(-x-y+1))}{-\delta + \varepsilon}\right) + \exp\left(\frac{6(y\delta - \varepsilon(-x-y+1))}{-\delta + \varepsilon}\right)}
\end{align*}
\]

Next, applying formula (2.10) for the computation of the first Lyapunov coefficient, we obtain after some further simplifications:

\[
l_1(\varepsilon, \delta) = \left[\frac{1728\delta\varepsilon - 4320\delta^2 - 4320\varepsilon^2 - 4320\delta\varepsilon + 1728\delta^2 + 1728\varepsilon^2}{19\delta^2 - 38\delta\varepsilon + 19\varepsilon^2 - 16\delta\varepsilon + 8\delta^2 + 8\varepsilon^2}\right]
= \frac{-2592\delta\varepsilon - 2592\delta^2 - 2592\varepsilon^2}{27\delta^2 - 54\delta\varepsilon + 27\varepsilon^2}
= -2592\frac{\delta\varepsilon + \delta^2 + \varepsilon^2}{3(3\varepsilon - 3\delta)^2}
\]
Chapter 3

Heterogenous Learning Rules in Cournot Games

3.1 Introduction

Stability of the Cournot equilibrium in homogenous oligopoly has been constantly scrutinized: from the seminal work of Theocharis (1960) who showed that the Cournot-Nash equilibrium of the standard linear inverse demand-linear cost model, loses stability once we allow for more than two oligopolists, through the ‘exotic’ chaotic dynamics discovered by Rand (1978) in specific non-linear Cournot models and to the relatively recent work on ‘evolutionary’ Cournot models be it in a stochastic (Alos-Ferrer (2004)) or deterministic framework (Droste et al. (2002)). Sources of instability vary from non-monotonicity in the reaction curves (Kopel (1996)) to alternative specification of the quantity adjustment or expectations-formation processes (Szidarovszky et al. (2000)). While it may not appear too surprising to generate complicated behaviour when one allows for ‘strange’ reaction functions, the sensitivity of the (linear) Cournot game outcome to the learning dynamics or to the interplay of different learning ‘heuristics’ seems more intriguing given also the growing body
of experimental evidence on the heterogeneity of such learning rules\(^1\).

As our goal is to explore the impact of such learning rules heterogeneity on the “limiting” outcome in Cournot oligopolies, we will review a few existing results concerning adjustment processes in oligopoly either in a static framework (two boundedly rational players endowed with Cournot expectations playing a best-reply to each other’s expectations) or in a population game (large group of oligopolists, randomly paired, with the resulting ecology of learning rules determined endogenously). Adaptive expectations are a straightforward generalization of Cournot expectations. They have been studied by Szidarovszky et al. (1994) who found that, in a multi-player oligopoly game, if the expectation adjustment speed is sufficiently low then the unique Cournot-Nash equilibrium is stabilized. Deschamp (1975) and Thorlund-Petersen (1990) investigate Fictitious Play duopolists, i.e. players choosing a quantity which is a best-reply to the empirical frequency distribution of the opponent past choices. Under certain restrictions on the cost structure (marginal costs not decreasing too rapidly) the process is shown to generate stable convergence to the Cournot solution. Milgrom and Roberts (1991) prove convergence for a wide range of adaptive processes provided that the oligopoly game is of so-called Type I - as coined by Cox and Walker (1998), and meaning that the reaction curves cross in such a way that the interior Cournot-Nash equilibrium is unique, i.e. there are no additional ‘boundary’ equilibria; this has to do, again, with the ‘marginal cost not decreasing too fast’ condition in Thorlund-Petersen (1990).

Turning to the ‘evolutionary’ type of modelling oligopoly, Vega-Redondo (1997) proved that imitation with experimentation dynamics leads away from the Cournot Nash equilibrium in a finite population of Cournot oligopolists. The proof uses techniques from perturbed Markov chain theory introduced to game theory by Kandori et al. (1993) to show that the Walrasian equilibrium is the only stochastically stable

---

\(^1\)See Camerer (2003) for a broad overview of the experimental games literature, Cheung and Friedman (1997) for experiments on learning in specific normal form games and Huck et al. (2002) for experimental evidence on learning dynamics in Cournot games.
state when the noise parameter goes to zero. (Alos-Ferrer (2004)) shows, in a similar stochastic setting, that Vega-Redondo (1997) result is not robust to adding finite memory to the imitation rule: anything between Cournot and Walras is possible as ‘long-run’ outcomes with only one period memory. Last, Droste et al. (2002) consider, in a deterministic set-up, evolutionary competition between ’Cournot’ players (endowed with freely available naive expectations) and rational or ‘Nash’ players (who can perfectly predict the choice of the opponent but incur a cost for information gathering). The resulting evolutionary dynamics is governed by the replicator dynamics with noise and leads to local and global bifurcations of equilibria and even strange attractors for some parameter constellations.

In this chapter, we construct an evolutionary oligopoly model where players are endowed with heterogenous learning procedures about the opponent’s expected behavior and where they update these routines according to a logistic evolutionary dynamic. In contrast with Droste et al. (2002), our focus is the interplay between adaptive and ‘fictitious play’ expectations, but other ecologies of rules (e.g. weighted fictitious play vs. rational or Nash expectations) are explored, as well. The choice of the Logit updating mechanism is motivated by the fact that this perturbed version of the Best-Reply dynamics, allows for an imperfect switching to the myopic best reply to the existing strategies distribution. The main finding is that in a population of fictitious (i.e. best-responders to the empirical distribution of past history of the opponent’s choices) and adaptive players (i.e. those who place large weight on the last period choice and heavily discard more remote past observations) the Cournot-Nash equilibrium can be destabilized and complicated dynamics arise. The chapter is organized as follows: Sections 3.2 and 3.3 briefly revisit the original Cournot analysis and the set of learning rules agents are endowed with. Section 3.4 introduces the model of evolutionary Cournot duopoly and evaluates alternative heuristics ecologies. Some concluding remarks are presented in section 3.5.
3.2 Standard Static Cournot Analysis

We consider a linear Cournot duopoly model much in line with the specification used in, for example, Cox and Walker (1998) or Droste et al. (2002). Assume a linear inverse demand function:

\[ P = a - b(q_1 + q_2), \quad a, b \geq 0, q_1 + q_2 \leq a/b, \]  

(3.1)

and quadratic concave costs (i.e.decreasing marginal cost),

\[ C_i(q_i) = cq_i - \frac{d}{2}q_i^2, \quad c, d \geq 0. \]  

(3.2)

In a standard maximization problem firm I chooses \( q_1 \) such that it maximizes instantaneous profits taking as given the quantity choice of its competitor \( q_2 \):

\[ q_1^* = \arg \max_{q_1} (Pq_1 - C_1) \]
\[ = \arg \max_{q_1} [(a - b(q_1 + q_2))q_1 - (cq_1 - \frac{d}{2}q_1^2)] \]  

(3.3)

and similarly for firm II,

\[ q_2^* = \arg \max_{q_2} [(a - b(q_1 + q_2))q_2 - (cq_2 - \frac{d}{2}q_2^2)]. \]  

(3.4)

This gives firm I and II’s reaction curves as:

\[ q_1 = R_1(q_2) = \frac{a - c - bq_2}{2b - d}; \quad q_2 = R_2(q_1) = \frac{a - c - bq_1}{2b - d} \]  

(3.5)

and their intersection yields the interior Cournot-Nash equilibrium of the game:

\[ q_1^* = q_2^* = \frac{a - c}{3b - d} \]  

(3.6)
The question of how equilibrium is reached and, in particular, how could one of the duopolists know the strategic choice of the opponent at the moment when she is making the quantity choice decision generated much debate in the literature. Cournot (1838) was already pointing to an equilibration process by which firms best-respond to the opponent’s choice in the previous period (in current terminology ‘naive’ expectations). However, there is a vast body of literature on whether and under what restrictions - for instance, the number of players and demand and cost function specifications - the Cournot process converges.

3.3 Heterogenous Learning Rules

3.3.1 Adaptive Expectations

In a boundedly rational environment, one possible reading of the reaction curves (3.5) is that each player best-responds to expectations about the other player’s strategic choice.

\[ q_1(t+1) = R_1(q_2^e(t+1)) \]
\[ q_2(t+1) = R_2(q_1^e(t+1)). \] 

A number of ‘expectation formation’ or ‘learning’ rules are encountered in the literature. Among them, particularly appealing are adaptive expectations where current expectations about a variable adapt to past realizations of that variable. Formally, expectations \( q_2^e(t+1) \) about the current period opponent quantity \(-q_2(t+1)\) are determined as a weighted average of last period’s expectations and last period actual choice.

\[ q_1^e(t+1) = \alpha_1 q_1^e(t) + (1-\alpha_1)R_1(q_2^e(t)) \] 
\[ = \alpha_1 q_1^e(t) + (1-\alpha_1)q_1(t) \]
\[ q_2^e(t + 1) = \alpha_2 q_2^e(t) + (1 - \alpha_2) R_2(q_1^e(t)) \] (3.9)

\[ = \alpha_2 q_2^e(t) + (1 - \alpha_2) q_2(t) \]

### 3.3.2 Fictitious Play

Fictitious play, introduced originally as an algorithm for computing equilibrium in games (Brown (1951)), asserts that each player best-responds to the empirical distribution of the opponent past record of play. In the context of Cournot ‘quantity-setting’ games this boils down to the average of the opponent past ‘played’ quantities where equal weight is attached to each past observation:

\[ q_2^e(t + 1) = \frac{1}{t} \sum_{k=1}^{t} q_2(k) \]

or FP given recursively:

\[ q_2^e(t + 1) = \frac{t - 1}{t} q_2^e(t) + \frac{1}{t} q_2(t), t \geq 1 \] (3.10)

### 3.3.3 Weighted Fictitious Play

Standard fictitious play could be generalized to allow for discarding the remote past observations at a higher rate. Cheung and Friedman (1997) use an exponentially-weighted scheme:

\[ q_2^e(t + 1) = \frac{q_2(t) + \sum_{u=1}^{t-1} \gamma^u q_2(t - u)}{1 + \sum_{u=1}^{t-1} \gamma^u} \]

or, recursively:

\[ q_2^e(t + 1) = \begin{cases} \frac{\gamma - \gamma^t}{1 - \gamma^t} q_2^e(t) + \frac{1 - \gamma}{1 - \gamma^t} q_2(t), \gamma \in [0, 1) \\ \gamma \frac{t-1}{t} q_2^e(t) + \frac{1}{t} q_2(t), \gamma = 1 \end{cases} \] (3.11)
It can be easily verified that Weighted Fictitious Play expectations nest, as special
cases, the following types of expectations:

(i) $\gamma = 0$: Cournot or ‘naive’ expectations/short memory;
(ii) $0 < \gamma < 1$: adaptive expectations with time-varying weights/intermediate
memory;
(iii) $\gamma = 1$: Fictitious Play/long memory.

### 3.4 Evolutionary Cournot Games

In this section we will consider an evolutionary version of the "static" model out-
lined in Section 3.2. Basically, at each time instance, two players are drawn randomly
from an infinite population and matched to play a standard duopoly quantity-setting
game. In order to form expectations about the other player’s quantity choice a set of
learning rules is available. Players switch between different rules based on some per-
formance measure. Rules that perform relatively better are more likely to spread in
the population and thus, fractions using each rule are expected to evolve over time
according to a specified updating mechanism. A more sophisticated expectation-
formation rule will come at a cost compared to the ‘simple’ rules which are assumed
costless. Furthermore, this heterogeneity of the learning process adds uncertainty
over the players expectation formation process. Since opponents may be of differ-
ent types, expectations are formed about an average opponent strategy or quantity
choice. A variety of such pairwise interactions and resulting evolution of the learning
rules is discussed below, analytically (when possible) and via simulations.

#### 3.4.1 Adaptive Expectations vs. Rational/Nash play

We start with a generalization of Droste et al. (2002) behaviors’ ecology by
considering a mixture of adaptive (nesting a special case of naive expectations as in
Droste et al. (2002)) and rational or ‘Nash’ players. The adaptive players behave
according to (3.8). Rational firms are able to correctly compute the quantity picked by an adaptive firm as well as the updated fractions of Adaptive and Rational players: their reaction function is playing a best-reply to the mixture of rules (equations (3.12), (3.13)):

\[
q_1(t) = R(q^e_1(t)) = \frac{a - c - bq^e_1(t)}{2b - d} \\
q_2(t) = (1 - n(t))q_1(t) + n(t)q_2(t) \tag{3.12} \\
q_2(t) = R(q^e_2(t)) = \frac{a - c - bq^e_2(t)}{2b - d} \tag{3.13} \\
\Rightarrow q_2(t) = \frac{a - c - b(1 - n(t))R(q^e_1(t))}{2b - d + bn(t)} \tag{3.14}
\]

where, \(n(t)\) is the fraction of type II, rational players and \(q^e_1(t), q^e_2(t)\) are the expectations about a randomly encountered opponent, of the adaptive and rational types, respectively.

The fraction of rational players \(n(t)\) updates according to the logistic mechanism with asynchronous updating (Hommes et al. (2005)). The asynchronous logistic updating has two components: an inertia component, parameterized by \(\delta\), showing the fraction of players who do not (are not given the opportunity to) update and the logistic probability of updating for the fraction \(1 - \delta\) of players for which a revision opportunity arises:

\[
n(t + 1) = \delta n(t) + \frac{(1 - \delta) \exp(\beta(U_2(q_2(t)) - k))}{\exp(\beta(U_1(q_1(t)) + \exp(\beta(U_2(q_2(t)) - k)))} \tag{3.15}
\]

with \(U_1, U_2\) representing the average profits of an adaptive and rational player at time \(t\), while \(\beta\) denotes the inverse of the noise/or ‘intensity of choice’ parameter and \(k\) stands for the extra-costs incurred by the agents employing the more sophisticated (here rational expectations) heuristic. Each strategy performance measure \(U_1(U_2)\) is computed as a fractions-weighted average of profits accrued in encounters with own
and other expectation-formation types:

\[ U_1 = (a - c)q_1 + \left( \frac{1}{2}d - b \right) q_1^2 - b(nq_1 + (1 - n)q_2)q_1 \]  
\[ U_2 = (a - c)q_2 + \left( \frac{1}{2}d - b \right) q_2^2 - b(nq_1 + (1 - n)q_2)q_2 \]  

Equations (3.15), (3.18) and (3.19) define a three-dimensional, discrete dynamical system:

The Cournot-Nash equilibrium \( q_1^*(t), q_2^*(t) \) for the adaptive and rational type, respectively - about the quantity of an average opponent quantity, evolve according to:

\[ q_1^*(t + 1) = \alpha q_1^*(t) + (1 - \alpha)((1 - n(t))q_1(t) + n(t)q_2(t)) \]  
\[ q_2^*(t + 1) = (1 - n(t + 1))q_1(t + 1) + n(t + 1)q_2(t + 1) \]  

Equations (3.15), (3.18) and (3.19) define a three-dimensional, discrete dynamical system \((q_1^*(t + 1), q_2^*(t + 1), n(t + 1)) = \Phi(q_1^*(t), q_2^*(t), n(t))\) describing both the evolution of the adaptive and rational types expectations and of the learning heuristics’ ecology:

\[ q_1^*(t + 1) = \alpha q_1^*(t) + (1 - \alpha)(n(t)R(q_1^*(t)) + (1 - n(t))R(q_2^*(t))) \]  
\[ q_2^*(t + 1) = \gamma q_1^*(t) + (1 - \gamma)(n(t)R(q_1^*(t)) + (1 - n(t))R(q_2^*(t))) \]  
\[ n(t + 1) = \frac{1}{1 + \exp(\beta(U_2(R(q_1^*(t)), R(q_2^*(t)), n(t)) - U_1(R(q_1^*(t)), R(q_2^*(t)), n(t)) - k))} \]

Using (3.12), (3.13) we can express firm/type II expectations as a function of firm I expectations about opponent quantity choice, effectively reduce (3.20) to a two-dimensional system:

\[ q_1^*(t + 1) = \alpha q_1^*(t) + (1 - \alpha)((1 - n(t))R(q_1^*(t)) + n(t)\frac{a - c - b(1 - n(t))R(q_1^*(t))}{2b - d + bn(t)}) \]  
\[ n(t + 1) = \delta n(t) + (1 - \delta)\frac{e^{(\beta(U_2(R(q_1^*(t)), n(t)) - k))} + e^{(\beta(U_2(R(q_1^*(t)), n(t)) - k))}}{e^{(\beta(U_2(R(q_1^*(t)), n(t)) - k))} + e^{(\beta(U_2(R(q_1^*(t)), n(t)) - k))}} \]

**Lemma 4** The Cournot-Nash equilibrium \( q_1^* = q_2^* = \frac{a - c}{3b - d} \) in (3.6) together with
\( n^* = \frac{1}{1+\exp(\beta k)} \) is the unique, interior, fixed point of (3.21). For low values of the intensity of choice \( \beta < \beta^* = \ln \left( \frac{(d-3b)(\alpha+1)}{b-d+3b\alpha-d\alpha} \right) \), the Cournot-Nash equilibrium is stable; at \( \beta = \beta^* \) a period-doubling bifurcation occurs and a stable two-cycle is born. The 2-cycle loses stability via a secondary Neimark-Sacker bifurcation when \( \beta \) hits a second threshold \( \beta^{**} = \beta^* + \frac{1+\delta}{(1-\delta)(b-d+\alpha(3b-d))((\alpha+1)(1-\frac{2}{b-\alpha}))} \) with a limit cycle surrounding each point of the 2-cycle.

**Proof.** At the fixed point \( q_1^*(t+1) = q_1^*(t) \) the following should hold:

\[
\frac{a-c-b(1-n(t))R(q_1^*(t))}{2b-d+bn(t)} = R(q_1^*(t)) \Leftrightarrow R(q_1^*(t)) = \frac{a-c}{3b-d}. \]

But, \( R(x) = \frac{a-c}{3b-d} \) implies \( x = \frac{a-c}{3b-d} \) and, thus, \( q_1^*(t) = \frac{a-c}{3b-d} \) is the unique, interior\(^2\) rest point of (3.21) with corresponding fraction given by \( n^* = \frac{1}{1+\exp(\beta k)}. \)

It can be shown that system (3.21) may be rewritten in terms of deviations \( Q_t = q_1^*(t) - q_1^* \) from the Cournot-Nash equilibrium steady state as \( (Q(t+1), n(t)) = F(Q(t), n(t)) \) where:

\[
Q(t+1) = -Q(t) \frac{b - 3b\alpha + d\alpha - bn(t)}{2b - d + bn(t)} \quad (3.22)
\]

\[
n(t+1) = \delta n(t) + \frac{1 - \delta}{1 + \exp(\beta(k - \frac{(Q(t)b(3b-d))}{2b-d+bn(t)}))} \]

This system has a fixed point at \( (Q^* = 0, n^* = \frac{1}{1+\exp(\beta k)}) \). In order to investigate it local stability we first derive the Jacobi matrix \( JF \) when evaluated at the steady state \( (Q^*, n^*) \) as:

\[
\begin{bmatrix}
-\frac{1}{2b-d+bn^*} (b - 3b\alpha + d\alpha - bn^*) & 0 \\
0 & \delta
\end{bmatrix}.
\]

\(^2\)We note that, besides the interior Cournot-Nash equilibrium, which is main focus of this Chapter, the type II (Cox and Walker (1998)) Cournot game considered, has two additional boundary equilibria (two extra intersections of the reaction functions on the \( x \) and \( y \) axis), namely: \( A = (q_1^A, q_2^A) = (\frac{a-c}{2b-d}, 0) \) and \( B = (q_1^B, q_2^B) = (0, \frac{a-c}{2b-d}) \). One can show that, in a non-evolutionary environment with homogenous Cournot (naive) expectations, the Cournot-Nash equilibrium is unstable and all initial conditions (except for the 45\(^\circ\) line) converge to a 2-cycle formed by the two Nash equilibria \( (A,B) \). However, these two additional boundary equilibria do not form a 2-cycle of our dynamical system.
Its eigenvalues are \( \lambda_1 = -\frac{1}{2b-d+bn^*} (b - 3b\alpha + d\alpha - bn^*) \) and \( \lambda_2 = \delta \). We see that the fixed point loses stability via a primary, period-doubling, bifurcation when \( \lambda_1 \) hits the unit circle:

\[
\lambda_1 = -\frac{1}{2b-d+bn^*} (b - 3b\alpha + d\alpha - bn^*) = -1 \iff \beta = \beta^* = \frac{1}{k} \ln \left( \frac{d-b\alpha}{b-d+3b\alpha-d\alpha} \right).
\]

At the resulting symmetric 2-cycle \( \{(Q, \bar{n}), (-\bar{Q}, \bar{n})\} \) the following equalities hold:

\[
Q(t+1) = -Q(t) \frac{b-3b\alpha + d\alpha - bn^*}{2b - d + bn^*} \\
Q(t) = -Q(t+1) \frac{b-3b\alpha + d\alpha - bn^*}{2b - d + bn^*}.
\]

Thus, \( \frac{b-3b\alpha + d\alpha - bn^*}{2b - d + bn^*} = -1 \iff \bar{n} = \frac{d-b\alpha(3b-d)}{2b} \). Next, \( Q \) solves \( \bar{n} = \delta \bar{n} + (1 - \delta) \frac{1 + \exp(\beta(\gamma - (\frac{2b-d}{2b-d+bn^*})^2))}{2} \), yielding

\[
Q_\pm = \pm \frac{1}{\sqrt{b}} \sqrt{1 - \alpha} (2b - d) \sqrt{\frac{\beta}{2b-d} \left( \beta \ln(\alpha+1) \frac{-3b+d}{b-d+3b\alpha-d\alpha} \right)}
\]

After further simplifications, the 2-cycle \( \{(\bar{Q}, \bar{n}), (-\bar{Q}, \bar{n})\} \) reads:

\[
\begin{bmatrix}
\left\{ \frac{1}{b} \sqrt{\frac{\beta-\beta^*}{2b-d}}, \frac{d-b\alpha(3b-d)}{2b} \right\}, \\
\left\{ \frac{1}{b} \sqrt{\frac{\beta-\beta^*}{2b-d}}, \frac{d-b\alpha(3b-d)}{2b} \right\} 
\end{bmatrix} (3.24)
\]

To investigate the stability of \( \{(\bar{Q}, \bar{n}), (-\bar{Q}, \bar{n})\} \) we have to look at the Jacobian of the compound map \( G = F^2 \) evaluated at the two-cycle.

\[
JG(\bar{Q}, \bar{n}) = JF(\bar{Q}, \bar{n}) JF((\bar{Q}, \bar{n})) = JF(\bar{Q}, \bar{n}) JF(-\bar{Q}, \bar{n}). \quad (3.25)
\]

The Jacobian matrix of \( F \) is:

\[
JF(Q, n) = \begin{bmatrix}
g_{11} & g_{12} \\
g_{21} & g_{22}
\end{bmatrix}
\]

71
with entries,

\[ g_{11} = -\frac{1}{2b-d+bn} (b - 3b\alpha + d\alpha - bn) \]

\[ g_{12} = Qb \left( \alpha - 1 \right) \frac{d-3b}{(2b-d+bn)^2} \]

\[ g_{21} = \frac{2Qb^2 \beta(1-\delta)(d-3b)^2(b-\frac{1}{2}d)}{(2b-d+bn)^2} \exp \left( \beta \left( k-Q^2b^2(d-3b)^2 \frac{b-\frac{1}{2}d}{(2b-d+bn)^2} \right) \right) \]

\[ g_{22} = \delta + \frac{2Qb^2 \beta(\delta-1)(d-3b)^2(b-\frac{1}{2}d)}{(2b-d+bn)^2} \exp \left( \beta \left( k-Q^2b^2(d-3b)^2 \frac{b-\frac{1}{2}d}{(2b-d+bn)^2} \right) \right) \]

The stability analysis can be simplified by exploiting the symmetry of the system (3.22):

\[ G = F^2(\bar{Q}, \bar{n}) = F(-\bar{Q}, \bar{n}) = F(T(\bar{Q}, \bar{n})) \]

with the transformation matrix \( T \) given by:

\[
T = \begin{bmatrix}
-1 & 0 \\
0 & 1 \\
\end{bmatrix}
\]

Each point of the 2-cycle is a fixed point of the map \( G = FT \) and so, the 2-cycle of \( F \) is stable iff the fixed points of \( G \) are stable. The stability of the 2-cycle is governed by the Jacobian of \( FT \):

\[
JG(\bar{Q}, \bar{n}) = JF(-\bar{Q}, \bar{n})T
\]

It can be shown that, after further algebraic manipulations, the Jacobian takes the following form:

\[
JG(\bar{Q}, \bar{n}) = \begin{bmatrix}
-1 & -\frac{4}{d-3b} \sqrt{-\frac{1}{2b} k^{\frac{\beta-\beta^*}{d-2b}}} \\
\hat{j}_{21} & \hat{j}_{22} \\
\end{bmatrix}
\]

with \( \hat{j}_{21} = \frac{1}{b} \frac{\beta}{1-\alpha} \left( \alpha + 1 \right) \left( 1 - \delta \right) (3b-d) (2b-d) (b-d+3b\alpha-d\alpha) \sqrt{k^{\frac{\beta-\beta^*}{2\delta(2b-d)}}} \) and \( \hat{j}_{22} = \delta - (\alpha + 1) \left( 1 - \delta \right) (\beta - \beta^*) (b-d+3b\alpha-d\alpha) \)

The matrix has a pair of complex conjugates eigenvalues as lengthy expressions of all the model parameters. Still, the determinant takes a more tractable form:

\[
\det JG = (\alpha + 1) \left( 1 - \delta \right) (\beta - \beta^*) (b-d+3b\alpha-d\alpha) - \delta + \frac{2}{b} \frac{k(\alpha+1)(1-\delta)(\beta-\beta^*)(b-d+3b\alpha-d\alpha)}{\alpha-1}.
\]
The Neimark-Sacker bifurcation condition \((\text{det} \, JG = 1)\) yields the intensity of choice threshold \(\beta^{**}\) at which a second bifurcation arises, namely:

\[
\beta^{**} = \beta^* + \frac{1 + \delta}{(1 - \delta)(b - d + \alpha(3b - d))(\alpha + 1)(1 - \frac{2}{b(1 - \alpha)})}
\]

This completes the proof of Lemma 4.

Summing up, as the intensity of choice varies from low to high values the system transits from a unique steady state, through a 2-cycle to two limit cycles at \(\beta = \beta^{**}\) when it undergoes a Neimark-Sacker bifurcation and each of the 2-cycle branches becomes surrounded by a closed orbit as in Fig. 3.1d. Panels (a)-(b) illustrate the limiting behavior of quantities picked by the two players as they become more responsive to payoff differences, while Panel (c) displays the time evolution of the proportion of the costly rational (or Nash) expectation-formation heuristic. The intuition is similar to Brock and Hommes (1997): when the system is close to the Cournot-Nash equilibrium the simple, costless adaptive rule performs well enough and, due to the evolutionary switching mechanism, a larger number of players makes use of this cheap heuristic. However with a vast majority of agents using adaptive rule the CNE destabilizes and, as fluctuations of increasing amplitude arise, it pays off to switch to the more involved, though costly, behavior (i.e. Nash) pays-off. Once a large majority of agents switches to the rational rule fluctuations are stabilized and agents switch back to the costfree heuristic, re-initializing the cycle. Fig. 3.2a-b illustrates the primary (period-doubling) and secondary (Neimark-Sacker) bifurcations along with the "breaking of the invariant circle" route to chaos as the intensity of choice and degree of expectations adaptiveness are varied, respectively. Finally, Panels (c)-(d) show a strange attractor in the quantity-fraction of Nash players and quantity-quantity subspaces.
Figure 3.1: Cournot duopoly game with an ecology of Adaptive and Rational players. Panels (a)-(b): quantities evolution for different values of the intensity of choice ($\beta$). The evolution of Nash players fraction is shown in Panel (c). Panel (d) displays a limit cycle, consisting of two closed curves around the two points of the unstable 2-cycle, arising from the secondary Hopf bifurcation. Game parameters set to $a = 17, b = 0.8, c = 10, d = 1.1, k = 1, \alpha = 0.05, \delta = 0.1$. 
Figure 3.2: Cournot duopoly game with an ecology of Adaptive and Rational players. Bifurcation diagrams of the equilibrium quantity chosen by the Adaptive players ($q_1$) with respect to the intensity of choice ($\beta$) (Panel (a)) and degree of expectation adaptiveness ($\alpha$) in Panel (b). Game parameters set to $a = 17, b = 0.8, c = 10, d = 1.1, \delta = 0.1$. Panels (c), (d) display projection of a strange attractor on the quantity-fraction of rational players and quantity-quantity subspaces, respectively. A strange attractor is obtained for the following parameterization: $a = 17.54, b = 0.754, c = 10, d = 1.1, k = 1.5, \alpha = 0.1, \beta = 7.8, \delta = 0.5$. 

(a) Bifurcation diagram: $\beta - q_1, \alpha = 0.05$

(b) Bifurcation diagram: $\alpha - q_1, \beta = 15$

(c) Strange attractor, $q_1 - n$

(d) Strange attractor, $q_1 - q_2$
3.4.2 Adaptive vs. Exponentially Weighted Fictitious Play

As was already pointed out, players form expectations about the strategy of the average opponent given the existing heuristics ecology (frequencies of each learning type). In this subsection, the subset available is restricted to Adaptive Expectations and Weighted Fictitious Play as introduced in subsection 3.3. Thus, the rational rule is replaced by a long-memory rule that relies only on past observed quantity information with, possibly, different weighting schemes for each observation. Similar to the previous sections, the more sophisticated heuristic comes at a cost $k$. Considering that a WFP process gives rise to a non-autonomous dynamical system we will discuss analytically the limit as $t \to \infty$ when the system becomes autonomous. Let $q_{AE}^e(t+1)$ and $q_{WFP}^e(t+1)$ denote the expectations formed about an average encountered opponent by an AE-player and a WFP-player, respectively. The evolution of these two learning rules, in a Cournotian strategic environment is given by the following system of difference equations:

$$q_{AE}^e(t+1) = q_1^e(t+1) = \alpha q_1^e(t) + (1 - \alpha)q_{avg}\text{opp}(t)$$
$$q_{WFP}^e(t+1) = q_2^e(t+1) = \gamma q_1^e(t) + (1 - \gamma)q_{avg}\text{opp}(t)$$

with $\gamma > \alpha^3$. By letting $n(t)$ denote the fraction of agents employing the more computationally-involved heuristic II (WFP expectations in this case) this system

---

$^3$In subsection 3.3 we have seen that $\alpha$ close to 0 renders adaptive expectations, while $\gamma$ approaching 1 yields the standard fictitious play. These two parameters could also be interpreted as proxies for "memory": when they are close to the lower bound 0 the past is heavily discarded and only recent observations matter for constructing expectations about opponent choice, while as they are near the upper bound 1 past information becomes increasingly important for future choices. Thus naive expectations [$\alpha = 0$] is a short-memory rule, as the entire past, but last period, is discarded, while fictitious play [$\gamma = 1$] is a long-memory rule, as each past observation is equally important.
can be explicitly re-written as:

\[ q_1(t + 1) = \alpha q_1(t) + (1 - \alpha)((1 - n(t))q_1(t) + n(t)q_2(t)) \]  
\[ q_2(t + 1) = \gamma q_2(t) + (1 - \gamma)((1 - n(t))q_1(t) + n(t)q_2(t)) \]  

The actual strategic choices of an AE- and WFP-player are simply best-responses to their expectations:

\[ q_1(t) = R(q_1^e(t)) = \frac{a - c - bq_1^e(t)}{2b - d} \]  
\[ q_2(t) = R(q_2^e(t)) = \frac{a - c - bq_2^e(t)}{2b - d} \]  

The fraction of the more sophisticated WFP players evolves according to an asynchronous updating logistic mechanism (Hommes et al. (2005)):

\[ n(t + 1) = \delta n(t) + (1 - \delta) \frac{\exp(\beta(U_2(t) - k))}{\exp(\beta U_1(t)) + \exp(\beta(U_2(t) - k))}, \]  

where \( U_1, U_2 \) represent the average profits of an AE and WFP player at time \( t \), and \( \beta \) denotes the inverse of the noise/or "intensity of choice" parameter and \( k \) stands for the extra-costs incurred by the agents employing the more sophisticated (i.e. with a higher weight placed on past observations about opponent choices) WFP heuristic. The performance measures \( U_1, U_2 \) are determined as in equations (3.16)-(3.17) from subsection 3.4.1. Similarly, we derive a three-dimensional, discrete dynamical system \((q_1^e(t+1), q_2^e(t+1), n(t+1)) = \Phi(q_1^e(t), q_2^e(t), n(t))\) describing both the evolution of the adaptive and fictitious play type expectations and of the learning heuristics’ ecology:

\[ q_1^e(t + 1) = \alpha q_1^e(t) + (1 - \alpha)((1 - n(t))R(q_1^e(t)) + n(t)R(q_2^e(t))) \]  
\[ q_2^e(t + 1) = \gamma q_2^e(t) + (1 - \gamma)((1 - n(t))R(q_1^e(t)) + n(t)R(q_2^e(t))) \]  
\[ n(t + 1) = \delta n(t) + \frac{1 - \delta}{1 + e^{(\beta U_1(R(q_1^e(t)), R(q_2^e(t), n(t))) - U_2(R(q_1^e(t)), R(q_2^e(t)), n(t)) + k))}} \]  

77
Lemma 5 The (static) Cournot-Nash equilibrium quantities \((q_1^*, q_2^*)\) determined in (3.6) together with \(n^* = \frac{1}{1+\exp(\beta k)}\) is an (interior) fixed point of system (3.31).

Proof. At the Cournot-Nash Equilibrium (CNE) \(q_1^* = q_2^* = \frac{a-c}{2b-d}\) we have, by definition, \(R(q_1^*) = q_1^*\) and \(R(q_2^*) = q_2^*\). Substituting into the first two equations of (3.31) gives the required fixed point equality. Then, \(n^* = \frac{1}{1+\exp(\beta k)}\) is the solution of \(n(t+1) = n(t)\) when this equation is evaluated at the CNE. ■

3.4.3 Local stability analysis

It can be shown that the Jacobian matrix of (3.31) evaluated at the equilibrium \((q_1^*, q_2^*, n^*)\) takes the following form:

\[
J_{\Phi/SS} = \begin{bmatrix}
\alpha + (1 - \alpha)(1 - n^*)r & (1 - \alpha)n^*r & 0 \\
(1 - \gamma)(1 - n^*)r & \gamma + (1 - \gamma)n^*r & 0 \\
0 & 0 & \delta
\end{bmatrix}
\]

(3.32)

where \(r = \frac{\partial R(q_1^*)}{q_1^*} = -\frac{b}{2b-d}, r < -1\).

We know that under homogenous, naive (or adaptive with small \(\alpha\)) expectations the CNE is unstable with a two-cycle emerging at the boundary. On the other hand, fictitious play alone (or weighted fictitious play with large weights put in remote past observations) stabilizes the CNE. Once the interplay between the cheap adaptive and costly WFP heuristics is introduced we can prove the following result:

Proposition 6 A system with evolutionary switching between adaptive and weighted fictitious play expectations (i.e. \(0 < \alpha \ll \gamma < 1\)) is destabilized when the sensitivity to payoff difference between alternative heuristics reaches a threshold \(\beta^{PD} = \frac{1}{k} \ln \frac{\gamma(r-1)-(1+\alpha)(r+1)}{(\gamma+1)(r+1)+\alpha(1-r)}\) and a stable two-cycle bifurcates from the interior CN steady state \((q_1^* = q_2^* = \frac{a-c}{2b-d}, n^* = \frac{1}{1+\exp(\beta k)}\).

Proof. For an ecology of Adaptive and Weighted Fictitious Play rules - \(\alpha \in\)
$(0, 1), \gamma \in (0, 1), \gamma \gg \alpha = \text{the eigenvalues of the Jacobian matrix (3.32) are: } \lambda_{1, 2} = \\
\frac{1}{2}r + \frac{1}{2} \alpha + \frac{1}{2} \gamma - \frac{1}{2} r \alpha + \frac{1}{2} r \alpha n^* - \frac{1}{2} r \gamma n^* \pm \frac{1}{2} \sqrt{S}, \lambda_3 = \delta \\
\text{with } S = r^2(\alpha^2(n^*)^2 - 2r^2 \alpha^2 n^* + r^2 \alpha^2 - 2r^2 \alpha \gamma (n^*)^2 + 2r^2 \alpha \gamma n^* + 2r^2 \alpha n^* \\
-2r^2 \alpha + r^2 \gamma^2 (n^*)^2 - 2r^2 \gamma n^* + r^2 + 2r \alpha^2 n^* - 2r \alpha^2 + 2r \alpha \gamma - 4r \alpha n^* \\
+ 2r \alpha - 2r \gamma^2 n^* + 4r \gamma n^* - 2r \gamma + \alpha^2 - 2 \alpha \gamma + \gamma^2 \\

Next the period-doubling condition $\lambda_2 = -1$ yields:

$$n^*(\beta) = \frac{(1 + \gamma)[(1 + r) + \alpha(1 - r)]}{2r(\gamma - \alpha)}$$

$$\beta_{PD} = \frac{1}{k} \ln \frac{\gamma(r - 1) - (1 + \alpha)(r + 1)}{(\gamma + 1)(r + 1) + \alpha(1 - r)} \quad (3.33)$$

We can also derive the intensity of choice threshold at which instability arises for an ecology of Naive vs. Weighted Fictitious Play - $\alpha = 0, \gamma \in (0, 1)$ - as a special case of (3.33):

$$\beta_{PD} = \beta^* = \frac{1}{k} \ln \frac{r(\gamma - 1) - (1 + \gamma)}{(\gamma + 1)(r + 1)}$$

For an ecology of Fictitious vs. Adaptive Players - $\gamma = 1, \alpha \in (0, 1)$ - system (3.31) is non-autonomous and we will investigate this case in a subsequent section, numerically.

We notice that if there are no costs associated with the sophisticated predictor $WFP$ (i.e. $k = 0$) the system is stable for any value of the intensity of choice. This is so because what trigger switching into the destabilizing adaptive rule are the costs incurred when using $WFP$; thus, for $k = 0$ there is no incentive to switch away from $WFP$ even in tranquil or stable periods.

The 2-cycle, in deviations from the Cournot-Nash steady state - $((x, y, n), (X, Y, n))$
with \( X = -x, Y = -y \) - solves the following system of equations:

\[
\begin{align*}
X &= \alpha x + (1 - \alpha)((1 - n)\frac{a-c-bx}{2b-d} + n\frac{a-c-by}{2b-d}) - \frac{(a-c)(1-\alpha)}{2b-d} \\
Y &= \gamma y + (1 - \gamma)((1 - n)\frac{a-c-bx}{2b-d} + n\frac{a-c-by}{2b-d}) - \frac{(a-c)(1-\gamma)}{2b-d} \\
x &= \alpha x + (1 - \alpha)((1 - n)\frac{a-c-bx}{2b-d} + n\frac{a-c-by}{2b-d}) - \frac{(a-c)(1-\alpha)}{2b-d} \\
y &= \gamma y + (1 - \gamma)((1 - n)\frac{a-c-bx}{2b-d} + n\frac{a-c-by}{2b-d}) - \frac{(a-c)(1-\gamma)}{2b-d} \\
n &= \delta n + \frac{1-\delta}{\exp\left(\frac{1}{\delta} \frac{b^2}{(a-d)^2}\right) \exp\beta((x-y)(2a-2c-4bx-2by+dx+dy+2bnx-2bny+k))}
\end{align*}
\]  

Given that the two-cycle above satisfies \( X = -x, Y = -y \), we obtain, after some manipulations, that the ratio of the two expectation rules deviations from equilibrium prediction is given by:

\[
\frac{x}{y} = \frac{(1 - \alpha)/(1 + \alpha)}{(1 - \gamma)/(1 + \gamma)}
\]

For \( \alpha \ll \gamma \) (an adaptive vs. fictitious play ecology) we get \( x > y \). Thus, the fluctuations of the adaptive, short-memory rule exceed in amplitude the oscillations incurred by the long memory heuristic (see Fig. 3.3ab).

If we analyze the time evolution of the costly WFP fraction we observe a similar pattern as for the naive vs. Nash ecology discussed in the previous subsection. In tranquil times, when the dynamics is close to the Cournot-Nash equilibrium, the cheap adaptive rule outperforms the more involved and costly WFP rule. Thus, through the selection mechanism, almost the entire population becomes adaptive. However, a Cournot duopoly with adaptive(near naive) expectations is unstable and, as fluctuations emerge, conditions are created for the costly, more sophisticated WFP rule, to take over the population and re-stabilize the system.

We can also show, numerically, that the 2-cycle loses stability via a secondary Neimark-Sacker bifurcation with a quasi-periodic attractor consisting of two closed curves, arising around the unstable branches of the 2-cycle (Fig. 3.3d). Furthermore, the invariant curves break into a strange attractor if the intensity of choice \( \beta \) is increased even further. Bifurcations diagrams with respect to different behavioral
parameters together with the evolution of a strange attractor are reported in Figures 3.4 and 3.5.

(a) $\beta = 2$. Time series

(a) $\beta = 7$. Time series

(c) $\beta = 7$. Evolution of WFP share

(d) $\beta = 4.25$. Phase plot

Figure 3.3: Cournot duopoly game with an ecology of Adaptive and Weighted Fictitious Play behaviors. Periodic and chaotic time series of the quantities evolution for different values of the intensity of choice (Panels (a)-(b)); evolution of the WFP heuristic share in the population (Panel (c)). Panel (d) displays quasi-periodic attractor in the space of quantities chosen by the duopolists. Game parameters: $a = 17, b = 0.8, c = 10, d = 1.1, k = 1.1, \alpha = 0.1, \gamma = 0.9, \delta = 0.1$. 

81
Figure 3.4: Cournot duopoly game with an ecology of Adaptive and Weighted Fictitious Play behaviors. Panels (a)-(d): bifurcation diagrams of the equilibrium quantity picked by the Adaptive type ($q_1$) with respect to degree of expectation adaptiveness ($\alpha$), weighting parameter in the WFP rule ($\gamma$), intensity of choice ($\beta$) and costs associated with WFP ($e$), respectively. Game parameters: $a = 15, b = 0.7, c = 10, d = 1.1, k = 1.1, \alpha = 0.1, \gamma = 0.9, \beta = 12, \delta = 0.1.$
Strange attractors, $q_1 - q_2, \beta = 13.5, \delta = 0.5$

Strange attractors, $q_1 - n, \beta = 13.5, \delta = 0.5$

Strange attractors, $q_1 - q_2, \beta = 10.8, \delta = 0.1$

Strange attractors, $q_1 - n, \beta = 10.8, \delta = 0.1$

Figure 3.5: Cournot duopoly game with an ecology of Adaptive and Weighted Fictitious Play behaviors. Long-run strange attractors in the quantity-quantity and quantity-share of Adaptive players, for low (Panels (a)-(b)) and high (Panel (c)-(d)) degree of syncronicity ($\delta$) in the updating process. Game and behavioral parameters: $a = 15, b = 0.7, c = 10, d = 1.1, k = 1.1, \alpha = 0.1, \gamma = 0.9, \delta = 0.1$. 

83
3.4.4 Naive vs. Fictitious Play

In this subsection we report numerical simulations of the long-run behavior of the evolutionary competition between costless naive expectations and costly fictitious play for a linear inverse demand-quadratic costs quantity-setting duopoly. This ecology combines, in a sense, two limiting results from the previous subsections: Cournot or naive expectations - as $\alpha \to 0$ - along with standard fictitious play expectations - as $\gamma \to 1$. It is known that fictitious play converges in a homogenous population of fictitious players to Cournot-Nash equilibrium while naive expectations alone generate continuous oscillations. Hence, a natural question to ask is whether the presence of fictitious players could stabilize a population of naive players. Fig. (3.6) presents numerical simulations of the competition between the two heuristics and show persistent fluctuations of the naive players quantity around the Cournot-Nash equilibrium. The share of fictitious players (Fig. 3.6c) lingers close to extinction and spikes occasionally taking over almost the entire population. Similar to the previous two subsections, what is driving this results are the costs associated with the fictitious play rule: when the entire population follows the FP rule, the CNE is stabilized and so there is an incentive to switch to the cheaper rule, i.e. adaptive heuristic. As more and more players become adaptive the CNE loses stability, fluctuations emerge and, thus, they create an advantageous environment for long-memory players to step in. Depending on the intensity of selection (as captured by parameter $\beta$) the fluctuations could be periodic (Panel(a)) of even chaotic: see Panel (d) in Fig. 3.6 for numerical plot of the largest Lyapunov exponent. Fig. 3.7a-d depicts the evolution of such an attractor as the sensitivity to material payoffs $\beta$ varies. In order to escape the long transient of the fictitious play dynamics, the phase portraits are plotted after skipping 100,000 points of the series. While the fictitious player quantities stay within a small neighborhood the Cournot-Nash equilibrium, the strategies of the Cornout(naive) player display chaotic fluctuations of large amplitudes.
Figure 3.6: Cournot duopoly game with Naive and Fictitious Play Heuristics. Pattern of play evolution for naive and fictitious players, for different values of the intensity of choice in Panels (a) and (b), respectively. Panel (c) shows the time evolution of Fictitious Play fraction while Panel (d) provides numerical evidence for chaotic dynamics for large values of $\beta$. Game and behavioral parameters: $a = 17$, $b = 0.8$, $c = 10$, $d = 1.1$, $k = 1$, $\alpha = 0$, $\delta = 0.1$. 
Figure 3.7: Cournot duopoly game with Naive and Fictitious Play Heuristics. Panels (a)-(d): the evolution of the phase portrait into a strange attractor as the intensity of choice increases. Notice that the quantity picked by the fictitious player wanders around the Cournot-Nash equilibrium quantity while the output choice of the naive players exhibits more wide chaotic fluctuations around the same CNE. Game and behavioral parameters: $a = 17, b = 0.8, c = 10, d = 1.1, e = 1, \alpha = 0, \delta = 0.1$. 
3.5 Concluding Remarks

We studied evolutionary Cournot games with heterogeneous learning rules about opponent strategic quantity or price choice. The evolutionary selection of such 'learning' heuristics is driven by a logistic-type evolutionary dynamics. While our focus is on the interplay between adaptive and 'fictitious play' expectations, other rules ecologies (e.g. weighted fictitious play vs. rational or Nash expectations) are explored, as well. Analytical and simulation results about the behaviour of the corresponding evolutionary Cournot games, when various parameters of interest change, are presented. In a Cournotian setting, the main finding is that a population of (weighted) fictitious players (i.e best-responders to the empirical distribution of past history of opponent choices) could be destabilized away from the Cournot-Nash equilibrium play by the presence of the adaptive expectation formation rule (i.e. players who place large weight on the last period choice and heavily discard more remote past observations). The interaction between a costly weighted fictitious play and the cheap adaptive expectations yields complicated quantity dynamics. The key mechanism driving dynamics is the endogenous switching between learning rules based on realized fitness: The fictitious play drive the system near the Cournot-Nash equilibrium where the simple, adaptive rule performs relatively well and, due, to its cost advantage, can invade and take over the more sophisticated predictor. However, the adaptive rule alone destabilizes the interior equilibrium and, when far from the CNE it pays off to bear the costs of the more involved rule reinitializing the evolutionary cycle.
Chapter 4

On the Stability of the Cournot Solution: An Evolutionary Approach

4.1 Introduction

Within a linear cost-linear demand set-up, Theocharis (1960) shows that once the duopoly assumption is relaxed and the number of players increases, the Cournot Nash equilibrium loses stability and bounded oscillations arise already for triopoly quantity-setting games. For more than 3 players oscillations grow quickly unbounded but they are stabilized by the non-negativity price and demand constraint. Hahn (1962) puts forth an alternative proof of this results and derives stability conditions for asymmetric cost functions and nonlinear cost and demand functions. Fisher (1961) and McManus (1964) recover stability of the equilibrium when firms use, instead of the discrete-time best response dynamic of Theocharis (1960), a continuous adjustment process to some optimally derived output target. However, firms do these adjustments incompletely (i.e. with different speeds of adjustment) at each time instance. In this class of adjustment processes, increasing the marginal costs acts as
a stabilizer of the Cournot-Nash equilibrium, irrespective of the number of players. In Okuguchi (1970) both the actual output and expectations of rivals’ output follow a continuous adjustment process to some target and restrictions on the speeds of adjustment (of both actual and the rival’s expected output) are derived in relation to the rate of marginal costs increase such that stability is regained for an arbitrary number of players. Adaptive expectations generalize Cournot (naive) expectations. Thus, in a multi-player oligopoly game, where players follow a discrete-time best-response to adaptively formed expectations about rivals output, Szidarovszky et al. (1994) find suitable restrictions on the coefficients of adaptive expectations such that the unique Cournot-Nash equilibrium is stabilized.

In this short Chapter we take a different route and relax the assumption of homogeneous expectations, while preserving the linear structure and best-reply adjustment dynamics of Theocharis (1960). The questions we address is whether Theocharis (1960) classical instability result still persist under the evolutionary selection of heterogeneous expectations rule. We show that, in an evolutionary environment, the instability is robust to a heterogenous ecology of heuristics. In particular, we focus on the interplay between a simple adaptive rule and a more involved, but costly, rational rule with agents choosing the best-performing rule according to a logistic updating mechanism. The threshold number of players that triggers instability may vary with the costs of the rational expectations rule or with the adaptive expectations coefficient.

In Section 4.2 the model is introduced, results and simulations are reported in Section 4.3 while Section 4.4 contains some concluding remarks.
4.2 The model

We start with the Theocharis (1960) simple linear inverse demand-linear costs homogenous, $n$-player oligopoly game. Linear inverse demand is given by

$$\mathcal{P} = a - b \sum_{i=1}^{n} q_i, \quad (4.1)$$

and linear costs by

$$C_i(q_i) = c q_i. \quad (4.2)$$

Each period $n$ players are drawn from a large populations to play a $n$-player quantity-setting game. Furthermore, assume that the population of oligopolists consists of two behavioural types: adaptive players (share $1 - p$ in the population) playing quantity $q_1$ and rational (share $p$) picking quantity $q_2$. From the perspective of an individual player the aggregate output of the remaining $n-1$ opponents, given behavioral types, could be computed for different market structures as follows:

In the triopoly case the two randomly drawn opponents may be, with probability $p^2$ of rational type (producing $q_2$ each), with probability $(1 - p)^2$ of adaptive type (and thus choosing $q_1$ each) and, with probability $2p(1 - p)$ of different types (thus picking $q_1$ and $q_2$, respectively). Thus, the output of an average opponent in a triopoly game is given by:

$$p^2 2q_2 + 2p(1-p)(q_2 + q_1) + (1 - p)^2 2q_1 \quad (4.3)$$

Similarly, for the quadropoly Cournot game we get

$$p^3 3q_2 + 3p^2(1-p)(2q_2 + q_1) + 3p(1-p)^2(q_2 + 2q_1) + (1 - p)^3 3q_1 \quad (4.4)$$
and for quintopoly:

\[ p^4q_2 + 4p^3(1-p)(3q_2 + q_1) + 6p^2(1-p)^2(2q_2 + 2q_1) + 4p(1-p)^3(3q_2 + 3q_1) + (1-p)^4q_1 \]

(4.5)

More generally, in the \( n \)-player quantity-setting game the output of an average opponent is given by:

\[
\sum_{k=0}^{n-1} \left[ \binom{n}{k} p^k (1-p)^{n-1-k} ((n-1-k)q_1 + kq_2) \right]
\]

\[
= q_2 \sum_{k=0}^{n-1} \left[ \binom{n}{k} p^k (1-p)^{n-1-k} \right] + q_1 \sum_{k=0}^{n-1} \left[ \binom{n}{k} p^k (1-p)^{n-1-k} (n-k-1) \right]
\]

\[
= q_2(n-1)p + q_1(n-1)(1-p)
\]

(4.6)

The aggregate expected output of the \( n-1 \) opponents is equivalent to the output of \( n-1 \) averaged-across-types opponents. Thus, each player has to form expectations about the output chosen by the average type in the population.

The adaptive type (AE) forms adaptive expectations about average the opponent:

\[ q_1^e(t+1) = \alpha q_1^e(t) + (1 - \alpha)[q_1(t)(1-p) + q_2(t)p]. \]

(4.7)

The rational type (RE) has perfect foresight:

\[ q_2^e(t+1) = q_1(t + 1)(1-p) + q_2(t + 1)p. \]

(4.8)

Both AE and RE types best-respond to expectations:

\[
q_1(t+1) = R_1(q_1^e(t+1)) = \frac{a-c-b(n-1)q_1^e(t+1)}{2b}
\]

\[
q_2(t+1) = R_2(q_2^e(t+1)) = \frac{a-c-b(n-1)q_2^e(t+1)}{2b}
\]

(4.9)
After substituting (4.9) and (4.11) into (4.7) we obtain the adaptive expectations dynamics as:

\[ q_1^e(t+1) = \alpha q_1^e(t) + (1-\alpha) [R_1(q_1^e(t))(1-p) + \frac{a - c + bR_1(q_1^e(t))(n-1)(1-p)}{2b(\frac{1}{2}p(n-1)+1)}] \]  

(4.12)

Next, we compute the expected profits as:

\[
\begin{align*}
\Pi_1(t) &= q_1(t)[a - b[q_1(t) + (n-1)[q_1(t)(1-p) + q_2(t)p]]] - c \\
\Pi_2(t) &= q_2(t)[a - b[q_2(t) + (n-1)[q_1(t)(1-p) + q_2(t)p]]] - c 
\end{align*}
\]

The share \( p \) of rational players updates according to asynchronous updating (only a fraction \( 1-\delta \) revise strategies according to the logistic probability) based on the realized performance measure:

\[ p(t+1) = \delta p(t) + \frac{1 - \delta}{1 + e^{\beta(\Pi_1(R(q_1^e(t)),R(q_2^e(t)),p(t)) - \Pi_2(R(q_1^e(t)),R(q_2^e(t)),p(t)) + k))}} \]  

(4.13)

where \( k \) stands for the costs associated with the rational expectations predictor (see the discussion regarding asynchronous logistic updating in Chapter 3). Equations (4.12) and (4.13) define a two dimensional dynamical system \((q_1^e(t+1), p(t+1)) = \Phi(q_1^e(t), p(t))\) which has an interior steady state at:

\[
(q_1^{e*} = \frac{a - c}{b(1+n)}, p^{*} = \frac{1}{1 + \exp(\beta k)})
\]
Expressed in terms of actual quantity choices the steady state is:

\[ q_1^* = q_2^* = \frac{a - c}{b(1 + n)}, \quad p^* = \frac{1}{1 + \exp(\beta k)} \]  

(4.14)

Before proceeding to the stability analysis of this equilibrium, we present in Fig. 4.1 simulations of the system behavior for varying number of players. As the game changes from triopoly to a large number of players the interior Cournot-Nash equilibrium (4.14) loses stability and bounded, regular or irregular oscillations arise. We identify a period-doubling route to strange attractors as the number of players increases.\(^1\) A snapshot of such a strange attractor is captured in Fig. 4.1e along with numerical evidence (strictly positive largest Lyapunov exponent) for chaos in Panel (f).

\[^1\text{This bifurcation diagram should be interpreted under the caveat that } n \text{ is not a continous parameter (the number of players in the Cournot game only takes positive integers value). Still, the bifurcation diagram can provide some information about qualitative changes in the behavior of the system, given the number of players.}\]
Figure 4.1: Linear $n$-player Cournot game with an ecology of Adaptive and Rational players. Panels (a)-(c) display converging, oscillating and chaotic time series of the quantity $q_1$ chosen by the Adaptive type for 3, 4 and 8-player game, respectively. Instability sets in already for the quadropoly game (Panel (b)). Panel (d) depicts a period-doubling route to chaotic dynamics in equilibrium quantity $q_1$ as we increase the number of players $n$. A typical phase portrait for an 8-player game is shown in Panel (e) while Panel (f) provides numerical evidence for chaos (i.e. positive largest Lyapunov exponent) as number of players increases. Game and behavioral parameters: $a = 17, b = 1, c = 10, k = 1, \alpha = 0.1, \beta = 2.8, \delta = 0.1$. 
4.3 Results

We start with the following:

**Proposition 7** The Cournot-Nash equilibrium \( q_1^* = \frac{a-c}{b(1+n)} \), \( p^* = \frac{1}{1+\exp(\beta k)} \) loses (local) stability as the number of player \( n \) reaches a critical threshold \( n_{PD} = \frac{2p^* - \alpha - 3}{2p^* + \alpha - 1} \) and a 2-cycle in the equilibrium quantities is born via a period-doubling bifurcation.

**Proof.** We can rewrite system in deviations from the Cournot-Nash steady-state

\[
Q(t) = \frac{a-c}{b(1+n)} - q_1^*(t)
\]

\[
Q(t + 1) = Q(t) \frac{1 + \alpha - (n + p(t)) + n(\alpha + p(t))}{p(t)(n - 1) + 2} \quad (4.15)
\]

\[
p(t + 1) = \frac{1 - \delta}{1 + \exp(\beta \left( \frac{-Q(t)(n^2 - 1) + (Q(t) - 2p) - 2cpn + 4c - Q(t) - 2np}{4n^2p^2 - 8np^2 + 16np^2 + 4p^2 - 16p + 16} + k \right))} \quad (4.16)
\]

with the interior steady state: \( (Q^* = 0, p^* = \frac{1}{1+\exp(\beta k)} \) ). The Jacobi matrix at this steady state is:

\[
\begin{bmatrix}
\frac{1}{p^*(n-1)+2} (\alpha - n - p^* + n (\alpha + p^*) + 1) & 0 \\
\frac{(4c - 2cn + 2cnp^*) \beta p^* (\delta - 1)(p^* - 1)((p^*)^2 - 1)}{4n^2(p^*)^2 - 8n^2p^* + 4n^2 + 16np^* - 16n + 16} & \delta
\end{bmatrix}
\]

with corresponding eigenvalues \( \lambda_1 = \frac{1}{np^* - p^* + 2} (\alpha - n - p^* + n\alpha + np^* + 1) \) and \( \lambda_2 = \delta \). The period-doubling bifurcation \( (\lambda_1 = -1) \) condition reads:

\[
\frac{1}{np^* - p^* + 2} (\alpha - n - p^* + n\alpha + np^* + 1) = -1,
\]

which yields,

\[
n_{PD} = \frac{2p^* - \alpha - 3}{2p^* + \alpha - 1}. \quad (4.17)
\]

We notice first that the Cournot-Nash equilibrium is stable if the rational predictor is costless. The threshold (4.17) can be re-arranged as \( \beta_{PD} = \frac{1}{k} \ln \left( \frac{(\alpha+1)(1+n)}{n-\alpha(1+n)+\delta} \right) \)
and observe that there is no finite value of the intensity of choice satisfying this equality for costless RE (i.e. \( k = 0 \)).

### 4.3.1 Best-response dynamics limit, \( \beta \to \infty \)

We can now derive Theocharis (1960) unstable triopoly result (i.e. bounded oscillations for \( n = 3 \) and exploding oscillations for \( n = 4 \)) in the limit of naive expectations \([\alpha \to 0]\) and best-response dynamics \([\beta \to \infty]\). First we compute \( \lim_{\beta \to \infty} p^* = 0 \) and then evaluate the period-doubling threshold (4.17) at \( \alpha \to 0 \) yielding

\[
\lim_{\beta \to \infty, \alpha \to 0} n^{PD}(p^*(k)) = 3.
\] (4.18)

The intuition behind the original "two is stable, three is unstable" result had to do with the slope of a reaction curves (4.9)

\[
\left| \frac{\partial R}{\partial q} \right| = \left| -\frac{n-1}{2} \right|
\]

For a duopoly game this slope is, in absolute values, always smaller than 1, but it hits the instability threshold exactly at \( n = 3 \) and is larger than 1 starting for a quadropoly game. This is exactly what our threshold \( n^{PD} \) boils down to when we let players choose a best reply \([\beta \to \infty]\) to Cournot (i.e. naively-determined) expectations \([\alpha \to 0]\).

Some further comparative statics can be informative about the role expectations play in (de) stabilizing a Cournot-Nash equilibrium. Ceteris paribus, the number of players’ instability threshold \( n^* \) increases in the degree of expectation adaptiveness \([\alpha]\):

\[
\frac{\partial n^{PD}}{\partial \alpha} = 4 \frac{1 - p^*}{(\alpha + 2p^* - 1)^2} > 0.
\]

Intuitively, a larger \( \alpha \) implies that the unstable adaptive rule place a higher weight on remote observations in the past and thus it acts as an effective stabilizing force. This, in turn, requires an even larger slope of the reaction curve (i.e. larger threshold \( n^* \)) to de-stabilize the interior CN equilibrium. Similarly, we can show that
\( n^* \) decreases in the costs associated with the perfect foresight rule:

\[
\partial n^{PD}(p^*(k))/\partial k < 0,
\]

as \( \partial n^{PD}(p^*)/\partial p^* = 4 \frac{\alpha + 1}{(2p + \alpha - 1)^2} > 0 \) and \( \partial p^*/\partial k < 0 \).

For a given number of players, at the steady state, higher costs of rational expectations \( k \), increase the payoff differential between the rational and the naive heuristics. Hence, more agents switch to the simple, yet destabilizing adaptive strategy. Thus, the system becomes more unstable and it takes a less steeper reaction curve (i.e. fewer players) to hit the period-doubling threshold (4.17).

### 4.3.2 Costly Rational Expectations, \( k > 0 \) and finite \( \beta \)

In the sequel we derive the threshold number of firms for which instability arises when there are positive costs associated with the perfect foresight and the other behavioural rule approaches either the Cournot (\( \alpha = 0 \)) or adaptive expectations (e.g. \( \alpha = 0.2 \)). Recall from (4.17) that \( n^{PD} = \frac{2p^* - \alpha - 3}{2p^* + \alpha - 1} \) and \( p^* = \frac{1}{1 + \exp(\beta \xi)} \). For small costs of the rational predictor (\( k = 1 \)) the period-doubling threshold is hit at:

\[
n^{PD}_{k=1} = \frac{\alpha + 3e^\beta + \alpha e^\beta + 1}{e^\beta - \alpha + \alpha e^\beta - 1} \tag{4.19}
\]

Table 4.1 reports unstable market structures in evolutionary Cournot games (i.e. number of players’ period-doubling, rounded thresholds (4.19)) for different adaptive expectations rule \( \alpha \) and intensity of choice \( \beta \).
These analytically derived thresholds are confirmed by the numerically-computed bifurcations curves illustrated Fig. 4.2a-d. Higher values of the intensity of choice may destabilize a given market structure: larger $\beta$ effectively implies turning on the evolutionary selection mechanism built into the model and population switching into/away from the adaptive, unstable heuristic. The two bottom panels in Fig. 4.2 display, for a quintopoly game, a strange attractors in the quantities chosen by the two types in the population (Panel (e)) and fractions of rational players-deviations from steady state quantity (Panel (f)).
Figure 4.2: Linear $n$-player Cournot game with costly Rational expectations. Bifurcation diagrams of the deviations $Q$ from equilibrium quantity $q_1$ with respect to number of firms $n$ for Naive [$\alpha = 0$] (Panel (a), (b)) and Adaptive [$\alpha = 0.2$] (Panel (c), (d)) expectations. Panels (a)-(d): The first period-doubling bifurcation threshold decreases in the intensity of choice parameter $\beta$. Panels (e)-(f) display projection of a strange attractor onto the quantity-quantity and deviations from CNE quantity-fractions of RE players, respectively. Game and behavioral parameters: $a = 17$, $b = 1$, $c = 10$, $k = 1$, $\delta = 0.1$. 

(a) $\alpha = 0$, $\beta = 1$. $n^{PD} =$Quintopoly  
(b) $\alpha = 0$, $\beta = 5$. $n^{PD} =$Triopoly  
(c) $\alpha = 0.2$, $\beta = 1$. $n^{PD} =$Decapoly  
(d) $\alpha = 0.2$, $\beta = 5$. $n^{PD} =$Quadropoly  

(e) $n = 5$, $\alpha = 0.2$, $\beta = 10.65$  
(f) $n = 5$, $\alpha = 0.2$, $\beta = 10.65$
4.4 Conclusions

We constructed an evolutionary version of Theocharis (1960) seminal work on the stability of the Cournotian equilibrium in multi-player quantity-setting games. The assumption of homogenous naive expectations is relaxed and players are allowed to choose between a costless, adaptive rule and a costly, rational rule. The results are consistent with the original homogenous-expectations analysis of Theocharis, but the instability thresholds vary with the costs of the sophisticated rule and the degree of adaptiveness in the expectation formation process. One implication of the model is that by fine-tuning these two parameters a particular Cournot market structure can be (de)stabilized. For instance, the classical unstable Cournot triopoly may be stabilized via making the rational expectations predictor freely available. On the other hand, when rational expectations are costly a period-doubling bifurcation route to chaos arises when the number of players increases.
Chapter 5

Evolution in Iterated Prisoner’s Dilemma Games under Smoothed Best-Reply Dynamics

5.1 Introduction

Studying evolutionary dynamics on iterated prisoner’s dilemma (IPD) demands a selection out of the vast set of the repeated game strategies. This choice turns out to be important for the outcome of the iterated game and the level of cooperation ‘evolved’. For instance, although the Tit-For-Tat strategy was an undisputed winner of two round-robin tournaments, Axelrod (1997) stresses the crucial role of the surrounding ecology of submitted iterated rules for the success of this direct reciprocity norm. However, which strategies to choose out of the large set of repeated PD is a question lacking a definite answer in the literature. Kraines and Kraines (2000) define adaptive dynamics\(^\text{1}\) on the space of memory one repeated PD strategies and show that this process selects, as long-run outcome, only three classes of repeated rules:

\(^{1}\text{Adaptive dynamics are regularly used in evolutionary biology to model evolution through natural selection on the fitness landscape.}\)
cooperative, alternating and defective. They also claim that, out of the cooperative subset, the Pavlov strategy is the only one that cannot be invaded by a strategy from the same 'cooperative' class (including TFT). Sigmund and Brandt (2006) investigate one such ecology consisting of three strategies: unconditional cooperators (AllC), unconditional defectors (AllD) and reactive players (TFT) and show that Replicator Dynamics exhibits, among others, a rock-scissors-paper pattern of cyclic behavior. However, as we have seen in Chapter 2, such replicator cycles are not robust under Replicator Dynamics, as small payoff perturbations drive all but one strategy near to extinction\(^2\).

In this chapter, we contribute to the stream of work on evolution of rules in repeated games by analyzing the repeated Prisoner’s Dilemma with a small number of simple, memory-one strategies. In particular, we extend Sigmund and Brandt (2006) ecology with two additional repeated strategies that seem to have received less attention in the evolutionary IPD game literature: the error-proof, "generous" tit-for-tat which, with a certain probability, re-establishes cooperation after a (possibly by mistake) defection of the opponent and the penitent, "stimulus-response" (thus dubbed Pavlov strategy Kraines and Kraines (1995), Sigmund and Nowak (1993b), Sigmund and Nowak (1993a)) that resets cooperation after the opponent punished for defection. Second, we contrast Replicator Dynamics behavior with a perturbed version of the Best-Reply dynamics, the Logit Dynamics, allowing for an imperfect switching towards a myopic best reply to the existing strategies distribution. Our emphasis is on the possibility of complicated dynamics such as multiplicity of steady states, limit cycles or chaos in the resulting dynamical system. We start with a systematic investigation of all ten \(2 \times 2\) repeated game rules interactions and then build-up towards more complex ecologies of 3, 4, and 5 strategies, combining theoretical considerations about the best-reply structure of the resulting normal form games with

\(^2\)mathematically, a so-called heteroclinic cycle is born, with Replicator Dynamics trajectories lingering longer and longer near the boundary of the simplex.
numerical analysis and simulations. A bifurcation analysis with respect to various model parameters is performed in order to reveal qualitative changes in the set of long-run (non)cooperative behaviors. Preliminary results, in particular for the \(4 \times 4\) and \(5 \times 5\) IPD ecologies, show that Logit Dynamics displays *stable coexistence* of repeated strategies but subjected to perpetual oscillations or even chaotic patterns in the distribution of IPD strategies.

The chapter is structured as follows: Section 5.2 introduces the selection of iterated PD strategies and the resulting evolutionary IPD game. The 2, 3 and 4-types ecologies are discussed in Sections 5.3, 5.4 and 5.5, respectively, while Section 5.6 investigates, mainly via computer simulations, the full \(5 \times 5\) ecology. The final section is reserved for concluding remarks and future research directions.

### 5.2 An Evolutionary Iterated PD game

We consider a standard \(2 \times 2\) Prisoner’s Dilemma stage game where players can either cooperate (C) or defect (D). The payoff matrix of the stage game is given by:

\[
\begin{bmatrix}
C/D & C & D \\
C & b - c, b - c & -c, b \\
D & b, -c & 0, 0
\end{bmatrix}; \quad b > c > 0
\]  

(5.1)

where \(b\) stands for the benefits of cooperation and \(c\) for the costs associated with cooperative behavior. At each time \(t\) the state of the play between two opponents is given by an element belonging to set \(\Omega = \{CC, CD, DC, DD\}\). For the iterated PD game we restrict the choice of iterated Prisoner’s Dilemma meta-rules to a set of stochastic *memory-one* strategies (see, for instance, Kraines and Kraines (2000)). Our motivation for such a simple subset of IPD strategies resides, on the one hand, on our attempt to model a particular form of boundedly rational players (namely limited memory agents or forgetting effects) and, on the other hand, to obtain analytical
tractability of the resulting Markov chain. The IPD game starts with a first random move C or D and then proceeds with playing C with probability \((r, s, t, p)\) conditional on the realized state at time \(t-1\) being \(CC, CD, DC, DD\), respectively. Deterministic strategies are particular limits in this stochastic strategy space. In particular we focus on 5 well-known strategies:

- unconditional cooperators \(AllC: (1,1,1,1)\);
- unconditional defectors \(AllD: (0,0,0,0)\);
- conditional cooperators "Tit-for-Tatters" \(TFT: (1,0,1,0)\);
- generous cooperators "Generous-Tit-for-Tat" \(GTFT: (1,m,1,n)\);\(^3\)
- penitent or Pavlov players "WinStayLoseShift" \(WSLS: (1,0,0,1)\).

Stochastic strategies are \(\varepsilon\)-perturbations of the deterministic ones, where \(\varepsilon\) has the natural interpretation of a probability of mistakes or errors in implementation/execution of the deterministic strategies. Following Kraines&Kraines (2000) the iterated Prisoner’s Dilemma game between two stochastic players \(S_1 = (r, s, t, p)\) and \(S_2 = (x, y, z, w)\) leads to a Markov chain on states \(CC, CD, DC, DD\) with transition probabilities given by:

\[
T = \begin{bmatrix}
    \text{state} & CC & CD & DC & DD \\
    CC & rx & sz & ty & pw \\
    CD & r(1-x) & s(1-z) & t(1-y) & p(1-w) \\
    DC & (1-r)x & (1-s)z & (1-t)y & (1-p)w \\
    DD & (1-r)(1-x) & (1-z)(1-s) & (1-t)(1-y) & (1-p)(1-w)
\end{bmatrix}
\tag{5.2}
\]

Each column of the matrix \(T\) in (5.2) contains transition probabilities to states \(CC, CD, DC, DD\), respectively and naturally, its entries add up to one. For instance,

\(^3m\) stands for the probability of cooperating after the opponent defected and \(n\) is the probability of playing cooperate after mutual defection.
in the first column, entry $t_{11}$ gives the probability of next period state staying in $CC$: row player $S_1$ cooperates with probability $r$ after the realized state was $CC$ while column player $S_2$ cooperates with probability $x$ after the same realization, and thus:

$$t_{11} = \text{prob}(CC | CC) = rx.$$  

Similarly,

$$t_{21} = \text{prob}(CD | CC) = \text{prob}(S_1 \text{ cooperates after } CC) \times \text{prob}(S_2 \text{ defects after } CC) = r \times (1 - \text{prob}(S_2 \text{ cooperates after } CC)) = r(1 - x).$$

One can show that, for strictly positive perturbation parameters, this Markov process is ergodic, i.e. there is positive probability of escaping from any of the states in $\Omega$. Therefore, it has a stationary invariant distribution given by the eigenvector of $T$ corresponding to an eigenvalue 1 (Kemeny and Snell (1975)). The invariant distribution represents the average time the play between two stochastic strategies spends in each state in $\Omega$ and, without discounting future payoffs$^4$, it enables computation of the average expected payoff resulting from the interaction of two repeated game strategies (Appendix 5.A contains detailed calculations of the invariant distributions and average payoffs). One such average payoff matrix is constructed below for an ecology consisting of the following stochastic versions of the five strategies described above:

- $AllC - (1 - \varepsilon, 1 - \varepsilon, 1 - \varepsilon, 1 - \varepsilon);
- AllD - (\varepsilon, \varepsilon, \varepsilon, \varepsilon);
- TFT - (1 - \varepsilon, \varepsilon, 1 - \varepsilon, \varepsilon);$

$^4$Sigmund and Brandt (2006) do allow for discounting in their 3x3 ecology of behaviors, and construct the iterated game matrix in a diferent manner, by computing discounted sums of future payoffs, under stochastic strategies; thus, the first move in the game becomes relevant and a repeated game strategy must be characterized, in addition, by the probability of playing C in the first round.
\begin{itemize}
  \item \textit{GTFT} \(- (1 - \varepsilon, m, 1 - \varepsilon, n^5)\);
  \item \textit{WSLS} \(- (1 - \varepsilon, \varepsilon, \varepsilon, 1 - \varepsilon)\).
\end{itemize}

\[
M = \begin{bmatrix}
  \text{AllD} & \text{TFT} & \text{GTFT} & \text{WSLS} & \text{AllC} \\
  \text{AllD} & \varepsilon (b - c) & m_{12} & m_{13} & \frac{1}{2}b - c\varepsilon & b - b\varepsilon - c\varepsilon \\
  \text{TFT} & m_{21} & \frac{1}{2}b - \frac{1}{2}c & m_{23} & \frac{1}{2}b - \frac{1}{2}c & m_{25} \\
  \text{GTFT} & m_{31} & m_{32} & \frac{n}{n + \varepsilon} (b - c) & m_{34} & m_{35} \\
  \text{WSLS} & b\varepsilon - \frac{1}{2}c & \frac{1}{2}b - \frac{1}{2}c & m_{43} & m_{44} & b - \frac{1}{2}c - b\varepsilon \\
  \text{AllC} & b\varepsilon - c + c\varepsilon & m_{52} & m_{53} & m_{54} & (1 - \varepsilon) (b - c)
\end{bmatrix} \quad (5.3)
\]

where \(m_{ij}\)'s are complicated algebraic expressions in the stage game parameters 
\((b, c)\), the errors \((\varepsilon)\) and the ‘generosity’ parameters \(n\) (see Appendix 5.A).

Players are assumed to switch their repeated game strategies based on realized, past \textit{average} performance. Thus, at time \(t\) the repeated strategy \(i\) will be played with the logistic probability:

\[
x_{i;t} = \frac{e^{\beta (Mx)_{i,t-1}}}{\sum_{i=1}^{5} e^{\beta (Mx)_{i,t-1}}}, \quad \sum_{i=1}^{5} x_{i,t-1} = 1 \quad (5.4)
\]

with \(\beta\) denoting the responsiveness to payoff differences between alternative strategies and \(x_{i,t}\)'s the time \(t\) fractions of \textit{ALLD, TFT, GTFT, WSL, ALLC} players in the population, respectively. Each entry in the payoff matrix (5.3) defines the \textit{long-run}\(^6\) contribution \(m_{ij}\) to the fitness of strategy \(S_i\) due to a particular encounter with strategy \(S_j\). This long-run contribution to fitness is used first to approximate the specific performance of strategy \(S_i\) in a \textit{finitely} repeated PD game and then to

\footnote{Throughout the Chapter we assume that the ‘generosity’ probabilities \(n\) and \(m\) are an order of magnitude higher that the probability of an error in the strategy execution \(\varepsilon\) (e.g. \(\varepsilon = 0.01, n = 0.1\)).}

\footnote{i.e. in the invariant distribution of the Markov chain}
update the behavior for the next IPD encounter.

In order to disentangle the rich behavior of the full 5-dimensional system (5.4) we start from simple pairwise interactions and then build larger ecologies of behaviors to capture the contribution of each type to the evolution of (non) cooperative outcomes.

## 5.3 2×2 Ecologies

In this section we discuss each of the \(\binom{5}{2} = 10\) cases of an ecology with only two repeated strategies. As we will see, depending on the stage game payoff matrix parameterization and of the error parameter \(\varepsilon\), the following interaction structures emerge for the IPD game: Coordination, Hawk-Dove, Prisoner’s Dilemma and dominance solvable games.

### 5.3.1 AllD vs. TFT

In a 2×2 ecology of unconditional defectors (AllD) and reciprocators (TFT) the reduced payoff matrix has the following form:

\[
\begin{bmatrix}
AllD & TFT \\
AllD & \varepsilon (b - c) & -\varepsilon (c - 2b + 2b\varepsilon) \\
TFT & \varepsilon (b - 2c + 2c\varepsilon) & \frac{1}{2}b - \frac{1}{2}c \\
\end{bmatrix}
\]  

(5.5)

For \(b > c/(1 - 2\varepsilon)\) and \(\varepsilon > 0\) the following inequalities hold:

\[
\varepsilon (b - c) > \varepsilon (b - 2c + 2c\varepsilon) \\
\frac{1}{2}b - \frac{1}{2}c > -\varepsilon (c - 2b + 2b\varepsilon)
\]
Thus, the IPD game (5.5) is a Coordination game\footnote{In the $\varepsilon \to 0$ limit TFT weakly dominates AllD and it is selected by the logit dynamics given when the sensitivity to payoffs difference is high enough (large $\beta$). However, taking this limit is problematic as the invariant distribution of the Markov chain is no longer unique, but history-dependent (more precisely the long-run payoff matrix now depends on the first move in the game). Still, for $\varepsilon = 0$, the payoff matrix could be regarded as the outcome of a particular unfolding of the game history.} with two pure strategy equilibria $(AllD, AllD)$ and $(TFT, TFT)$ and, depending on the initial population mixture, the Best-Reply limit $(\beta \to \infty)$ of the Logit Dynamics converges to either of them. When implementation errors are very small ($\varepsilon = 0.01$), the $(AllD, AllD)$ equilibrium basin of attraction is much smaller compared to the other equilibrium basin. This is due to the fact that $TFT$ against itself performs much better than $AllD$ against itself, as can be seen from the payoff matrix (5.5): $m_{22} = 1/2(b - c) > m_{11} = \varepsilon(b - c)$.

It turns out that, for large $\beta$ and small $\varepsilon$, $TFT$ can invade and take over the entire population, irrespective of the initial mixture (Fig. (5.1), Panel (a) show time series of the fractions for highly asymmetric starting conditions of 99% AllD and only 1% TFT). Increasing the probability of mistakes $\varepsilon$, increases the payoff of defectors against themselves and, in the $\beta$ large limit, the system displays co-existence of stable steady states. A certain large initial critical mass of defectors (fraction of AllD 99% and TFT 1%) can eliminate TFT with the population ending up in a monomorphic AllD state (Fig. 5.1b). The multiple steady states are created via saddle-node bifurcations: e.g. (Fig. 5.1cd) there is an unique interior fixed point for low values of $\beta(\varepsilon)$ but as the intensity of choice(mistake probability) reaches certain thresholds $\beta \approx 140(\varepsilon \approx 0.05)$ two additional steady states are created: one stable(both states in the panels) and one unstable. This behavior is similar to the evolutionary coordination game studied in Chapter 2 (see e.g. Fig. 2.10a).
Figure 5.1: Unconditional defectors (AllD) vs. Reciprocators (TFT). Different errors in the implementation lead to different long-run steady states: a population of reciprocators (Panel (a)) or a population of defectors (Panel (b)). Panels (c)-(d) display co-existence of stable steady states for large values of $\beta$ and $\varepsilon$. Remaining game parameters: $b = 4, c = 1, \beta = 200$.

5.3.2 TFT vs. AllC

For an interaction between unconditional cooperators (AllC) and reciprocators (TFT) the payoff matrix (5.3) reduces to:

$$
\begin{bmatrix}
TFT & \frac{1}{2}(b - c) & b - c - b\varepsilon + 2c\varepsilon - 2c\varepsilon^2 \\
AllC & b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2 & (1 - \varepsilon)(b - c)
\end{bmatrix}
$$

(5.6)
It is straightforward to see that, for small \( \varepsilon \) and \( b > c \) we have:

\[
\begin{align*}
    b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2 &> \frac{1}{2}(b - c) \\
    b - c - b\varepsilon + 2c\varepsilon - 2c\varepsilon^2 &> (1 - \varepsilon)(b - c)
\end{align*}
\]

Consequently, the reduced \( 2 \times 2 \) game (5.6) is of Hawk-Dove type and, thus, it has three equilibria: two asymmetric in pure strategies \((TFT, AllC)\) and \((AllC, TFT)\) and one symmetric in mixed strategies. The two asymmetric equilibria give rise to a 2-cycle under best-reply dynamics (i.e. logit dynamics with large \( \beta \)) while the interior equilibrium is unstable. A two-cycle, with population swinging back and forth between the two asymmetric equilibria, is created via a period-doubling bifurcation when either errors in implementation are small (Fig. 5.2c) or players choose a best-reply (large \( \beta \), Panel(d)). The intuition for the two-cycle relates to the best-response correspondences in the payoff matrix above: \( TFT \in BR\{AllC\} \) and \( AllC \in BR\{TFT\} \). In the population game interpretation, when everyone plays \( TFT(AllC) \) it is better to switch to \( AllC(TFT) \). Still, the amplitude of the cycle fluctuations varies with how responsive players are to payoff differences, with the full-scale two-cycle observed only for large \( \beta \).

For intermediated values of \( \varepsilon \) (e.g. \( \varepsilon \in \left( \frac{1}{4}, \frac{1}{2} \right) \) for the \( b = 2, c = 1 \) game parameterization) matrix (5.6) defines a dominance solvable (DS) game with \((TFT, TFT)\) as unique pure strategy equilibrium while for large \( \varepsilon > \frac{1}{2} \) the game is still dominance solvable, but with \((AllC, AllC)\) the unique equilibrium. The transition from a Hawk-Dove to a Dominance Solvable structure in the space of \( 2 \times 2 \) games is illustrated in Fig. 5.2ef for finite and best-reply limit of \( \beta \), respectively.
Figure 5.2: Reciprocators(TFT) vs Unconditional Cooperators(AllC). Time series of fractions of TFT and AllC players for small (Panel (a)) and large (Panel (b)) intensity of choice values, respectively. Panels (c)-(d) display a period-doubling bifurcation with respect to $\varepsilon$ and $\beta$, leading to a 2-cycle. Last, Panels (e)-(f) show transition from a Hawk-Dove to Dominance Solvable pattern. Game parameters: $b = 2, c = 1, \varepsilon = 0.01$. 
5.3.3 AllD vs. AllC

For a $2 \times 2$ ecology consisting of undiscriminating defective and cooperative types the restricted payoff matrix reads:

$$
\begin{bmatrix}
  AllD & AllC \\
  AllD & \varepsilon (b - c) & b - b\varepsilon - c\varepsilon \\
  AllC & b\varepsilon - c + c\varepsilon & (1 - \varepsilon) (b - c)
\end{bmatrix}
$$

(5.7)

We have:

$$
\begin{align*}
\varepsilon (b - c) & > b\varepsilon - c + c\varepsilon \\
 b - b\varepsilon - c\varepsilon & > (1 - \varepsilon) (b - c)
\end{align*}
$$

and the Iterated Prisoner’s Dilemma is a PD game itself with $(AllD, AllD)$ as the unique equilibrium. This case illustrates the standard prisoner’s dilemma argument: unless errors in implementation are extremely high (e.g. $\varepsilon \approx 0.5$ or $\beta$ very small, both making the players choice virtually random as in Fig. (5.3)a), the unconditional cooperators are heavily exploited by defectors, a situation leading to full extinction of the cooperative "trait" from the population (Fig. (5.3)b). The result is robust to any initial mixture of the population. For small finite $\beta$ or relatively large mistake probability there is a unique interior steady state of fractions $(x^*_{AllD}, x^*_{AllC})$ with $x^*_{AllD}$ sliding towards 1 as $\beta \to \infty$ and/or $\varepsilon \to 0$. Fig. (5.3)c-d illustrates the continuation of the equilibrium with respect to each of the two parameters.
Figure 5.3: Unconditional defectors (AllD) vs. Unconditional cooperators (AllC). Unless the choice is almost random (Panel (a), $\beta = 0.05$), the system quickly converges to a monomorphic state with defectors only (Panel (b)). Unique steady state converging to an AllD population when $\beta$ is large or $\varepsilon$ gets smaller (Panels (c)-(d)). Game parameters: $b = 2, c = 1, \varepsilon = 0.01$.

### 5.3.4 AllD vs. GTFT

For a sub-ecology of undiscriminating defectors and generous reciprocators the game matrix has the form:
\[
\begin{pmatrix}
  \text{AllD} & \text{GTFT} \\
  \varepsilon (b - c) & b\varepsilon - c\varepsilon - b\varepsilon^2 + bn - bn\varepsilon \\
  b\varepsilon - c\varepsilon + c\varepsilon^2 - cn + cn\varepsilon & \frac{n}{n+\varepsilon} (b - c)
\end{pmatrix}
\]  

(5.8)

which is always a coordination game for a ‘reasonable’ IPD parameterization.

This behaviors ecology generalizes the case discussed in subsection (5.3.1). Similarly, the system displays co-existence of monomorphic steady states for the $\beta \to \infty$ limit. Unlike the AllD vs. TFT case the multiplicity of steady states is preserved even for very small $\varepsilon = 0.01$ (Fig.5.4a-b) with AllD taking over the population provided that it slightly outnumbers TFT in the initial mixture (Fig.5.4a). The crucial parameter for this robust path-dependence, is the degree of generosity $n$ in GTFT strategy: "generous" TFT plays cooperate with quite large probability ($n = 0.3$) even after the opponent defected and this makes it more prone to exploitation from AllD compared to the typical TFT strategy. Reversing the initial mixture leads to long-run behaviour qualitatively similar to the one observed in the AllD vs. TFT sub-ecology (Fig.5.4b).

Panels (c)-(d) display the fold bifurcation scenarios through which the co-existing monomorphic steady states emerge. For instance, in Panel (d) as the critical threshold $n \approx 0.3$ is hit, a new stable fixed point is created with AllD players only for a parameterization ($\varepsilon = 0.01, \beta = 200$) for which uniqueness of the steady state was observed in the AllD vs. TFT ecology. By carefully tuning in the generosity parameter in the GTFT strategy, one could generate co-existing stable steady states in an AllD – (G)TFT ecology of repeated rules.

\footnote{AllD is a best reply to itself iff $\varepsilon^2 + n\varepsilon - n < 0$ (satisfied as long as $n$ is an order higher than $\varepsilon$); similarly, GTFT is best-response to itself as long as $\varepsilon^2 + n\varepsilon - n < 0$ and $b > c$.}
(a) Initial mixture: AllD-60%, GTFT-40%
(b) Initial mixture: AllD-40%, GTFT-60%

(c) $\varepsilon = 0.01, n = 0.3$. Bifurcation diagram ($AllD, \beta$)
(d) $\varepsilon = 0.01, \beta = 200$. Bifurcation diagram ($AllD, n$)

Figure 5.4: Unconditional defectors (AllD) vs. Generous reciprocators (GTFT). Long run behavior strongly depends on the initial mixture of the population even for very small errors in implementation. Panels (c)-(d): co-existence of monomorphic steady states(created via fold bifurcations) when $n$ and $\beta$ are large, even for tiny implementation errors $\varepsilon$. Baseline game parameters: $b = 2, c = 1, \varepsilon = 0.01, n = 0.3, \beta = 15$.

### 5.3.5 AllD-WSLS

In this sub-case, the payoff matrix (5.3) reduces to:

\[
\begin{bmatrix}
\text{AllD} & \varepsilon (b - c) & \frac{1}{2}b - c\varepsilon \\
\text{WSLS} & b\varepsilon - \frac{1}{2}c & (b - c) (1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon)
\end{bmatrix}
\] (5.9)
For high enough benefits of cooperation \((b > b^{*9})\), this payoff matrix represents a Coordination game if:

\[
\varepsilon (b - c) > b \varepsilon - \frac{1}{2} c
\]
\[
(b - c) (1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon) > \frac{1}{2} b - c \varepsilon
\]

In this case two pure strategy equilibria \((AllD, AllD)\) and \((WSLS, WSLS)\) exist. One could notice first that, for small implementation errors \(\varepsilon\), Pavlov fares much better against itself than AllD does against itself \(((b - c) (1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon) > \varepsilon (b - c))\). However, this may not be enough to offset the heavy exploitation Pavlov incurs in mixed encounters when AllD gets, on average, half from the benefits \(\lim_{\varepsilon \to 0} \frac{1}{2} b - c \varepsilon\) while WSLS pays half of the costs \(\lim_{\varepsilon \to 0} (b \varepsilon - \frac{1}{2} c)\). This situation is depicted in Fig. (5.5)a where the population ends up in an all-defectors state. If, in addition to small error probabilities, the benefits of cooperation are relatively large\((b > b^{*})\), the "own-type" interaction effect may dominate the "cross" interaction effect and the whole population becomes entirely Pavlovian (Fig. 5.5b). However, this results is dependent on the initial mixture of the population as for a given initial threshold of AllD players, population cannot turn Pavlovian, irrespective of the relative size of the benefits accrued to cooperation \(b\). The co-existence of two stable, monomorphic steady states for high benefits of cooperation is documented in Fig. 5.5c-d where additional fixed points emerge as the parameters of interest \(-\beta\) and \(\varepsilon\) respectively - are changed. When benefits are low enough, \(b < b^{*}\) the 2 x 2 game becomes dominance solvable (DS) with \((AllD, AllD)\) the unique equilibrium. A bifurcation diagram with respect to the benefits of cooperation \(b\) (Fig. 5.5e) reveals once the threshold \(b^{*}\) is passed, a transition in the space of 2 x 2 games from a Dominance Solvable game with unique equilibrium to a Coordination game with two equilibria.

\[b^{*} = 2c(1 - 2\varepsilon + 2\varepsilon^2)/(-4\varepsilon + 4\varepsilon^2 + 1)\]
Figure 5.5: Unconditional defectors (AllD) vs. Pavlov (WSLS). The selected long-run monomorphic state, depends both on the size of the implementation errors $\varepsilon$ and on the relative benefit of cooperation $b$. Panels (c)-(d): fold bifurcations in $\varepsilon, \beta$ spaces and co-existence of steady states. Panel (e) displays the bifurcation diagram with respect to the benefit of cooperation $b$ and show the transition from a dominance solvable game (low $b$) to a coordination game (high $b$). Panel (f) reports time series converging to an AllD population as WSLS is a strictly dominated strategy for $b < 2c$. Baseline game parameters: $b = 4, c = 1, \varepsilon = 0.01, \beta = 15$. 

(a) $b = 4$. Initial mixture (70%, 30%)
(b) $b = 4$. Initial mixture (50%, 50%)
(c) $\varepsilon = 0.01$. Bifurcation diagram
(d) $\beta = 200$. Bifurcation diagram
(e) $\beta = 200, \varepsilon = 0.01$. Bifurcation
(f) $b = 2$. Time series
5.3.6 GTFT vs. AllC

The $2 \times 2$ payoff matrix for an ecology of generous reciprocators and unconditional cooperators yields:

$$
\begin{bmatrix}
\text{GTFT} & \text{AllC} \\
\text{GTFT} & \frac{n}{n+\varepsilon} (b - c) \\
\text{AllC} & b - c - 2b\varepsilon + c\varepsilon + b^2\varepsilon + bn\varepsilon
\end{bmatrix}
$$

For $b > \frac{c}{1-n-\varepsilon}$\textsuperscript{10} and $n + \varepsilon < 1$, the following inequalities hold:

$$
b - c - 2b\varepsilon + c\varepsilon + b^2\varepsilon + bn\varepsilon \ > \ \frac{n}{n + \varepsilon} (b - c)
$$

$$
b - c - b\varepsilon + 2c\varepsilon - c^2\varepsilon - cn\varepsilon \ > \ (1 - \varepsilon) (b - c)
$$

They define a Hawk-Dove best-response structure: $\text{AllC} \in \text{BR}\{\text{GTFT}\}, \text{GTFT} \in \text{BR}\{\text{AllC}\}$ and, thus, generalize the $2 \times 2$ ecology $TFT$ vs. $AllC$ discussed in subsection 5.3.2. In a population of $GTFT$ players it pays off to be an unconditional cooperator because generous reciprocators ($GTFT$) do not exploit the opponent. On the other hand, in a population of AllC but subject to random errors, a strategy sensitive to those errors like $GTFT$ does better than an undiscriminating cooperator.

In contrast with the $TFT$ vs. $AllC$ ecology, the threshold at which $AllC$ invades $GTFT$ is always smaller than 1, i.e. there is always a fraction of $AllC$ players in the population (see. Fig. 5.6b-c). This fraction of permanent $AllC$ players goes up as the generosity parameter $n$ increases. This is due to the fact that with larger $n$, $GTFT - (1 - \varepsilon, n, 1 - \varepsilon, n)$ approaches $AllC - (1 - \varepsilon, 1 - \varepsilon, 1 - \varepsilon, 1 - \varepsilon)$ in the space of stochastic strategies, and, thus, it provides a more favorable environment for $AllC$.\footnote{If this condition is not satisfied, then the game is dominance solvable with a unique pure strategy equilibrium ($GTFT,GFT$).}

In the small $\beta$ case, and for small relative benefits of cooperation, the payoff differences between the two strategies vanish and we observe coexistence of behaviors
as long run outcome (Fig. 5.6a).

In sum, the long-run behavior of this ecology can lead to either a steady state with co-existing strategies - for large $n$ and $\beta$ (Panel (d)) or for small $b$ and $\beta$ (Panel(a)) - or to a stable two-cycle for small $n$ and large $\beta$ (Panels, (b)-(c)).

Figure 5.6: Generous reciprocators (GTFT) vs. Unconditional cooperators (AllC). Polymorphic population for relatively small benefit of cooperation and intensity of choice parameters (Panel (a)) and two-cycle for increased $\beta$ (Panel (b)) or $b$ (Panel (c)). Panel (d): a two-cycle created via period-doubling bifurcation with respect to the 'generosity' parameter $n$. Game parameters: $c = 1$, $\varepsilon = 0.01$, $n = 0.3$.

5.3.7 GTFT vs. WSLS

For this pair of repeated strategies the payoff matrix (5.3) reduces to:
\[
\begin{bmatrix}
\text{GTFT} & \frac{n}{n+\varepsilon} (b - c) & m_{34} \\
\text{WSLS} & m_{43} & (b - c) \left(1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon\right)
\end{bmatrix}
\] (5.11)

where \(m_{43}\) and \(m_{34}\) are complicated expressions of all model parameters (see Appendix (5.A) for explicit formulas)). One can show, by algebraic manipulations, that for \(n > \varepsilon\) (by assumption):

\[
\frac{n}{n + \varepsilon} (b - c) > m_{43}
\]

\[
(b - c) \left(1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon\right) > m_{34}
\]

and this \(2 \times 2\) ecology defines a Coordination game with two pure strategy, symmetric Nash equilibria (\(\text{GTFT, GTFT}\)) and (\(\text{WSLS, WSLS}\)). In the \(\beta\) large limit, the equilibria structure of the coordination game describe the long-run behavior of the logit dynamics. The corresponding basins of attraction relative sizes are significantly biased towards Pavlov players as it takes 70\% of \(\text{GTFT}\) players in the initial population in order to converge to a long-run all \(\text{GTFT}\) state (Fig.5.7a), with anything below this threshold reversing fortunes toward a long run Pavlov monomorphic state.

Intuitively, a stimulus-response, past-performance driven strategy like Pavlov, is able to take advantage of the generosity built into \(\text{GTFT}\) which reverts, with large probability \(n\), to "nice" behaviour after an opponent defection irrespective of the realized history. This is because \(\text{GTFT}\) does not discriminate among possible causes of defection (be it after a DD, or a DC history), while Pavlov re-establishes cooperation only after own defection. For finite, small \(\beta\) the population converges to a unique fixed point with co-existing behaviors of the form \((x^*_\text{GTFT}, x^*_\text{WSLS}), x^*_\text{GTFT} + x^*_\text{WSLS} = 1\), with Pavlov outnumbering generous reciprocators \((x^*_\text{WSLS} \text{ slides up in the interval } [1/2, 1] \text{ as we increase } \beta\)\). Fig. 5.7c-d depict a saddle-node bifurcation route towards co-existing monomorphic steady states. For low \(\beta\) (given \(n\)) or small \(n\) (fixing \(\beta\) close to the "best-reply" limit) there is a unique polymorphic population steady state.
Figure 5.7: Generous reciprocators (GTFT) vs. Pavlovian (WSLS). Depending on the initial fractions in the populations, in the $\beta$ large limit, the system converges to a monomorphic state with only one type of behavior surviving. Panels (c)-(d): creation of two co-existing steady states via a fold bifurcation when the degree of generosity/intensity of choice is varied. Game parameters: $b = 4, c = 1, \varepsilon = 0.01, n = 0.3, \beta = 100$. 

(a) Time series, initial mixture 70% GTFT-30% WSLS

(b) Time series, initial mixture 50% GTFT-50% WSLS

(c), $n = 0.3$. Bifurcation diagram ($WSLS, \beta$)

(d) $\beta = 100$. Bifurcation diagram ($WSLS, n$)
5.3.8 TFT vs. GTFT

For a population of generous and standard tit-for-tatters, the payoff matrix has the form:

\[
\begin{bmatrix}
TFT & \frac{1}{2}b - \frac{1}{2}c & GTFT \\
TFT & \frac{b\varepsilon - c\varepsilon^2 + bn - cn - 2bn\varepsilon + cn\varepsilon}{n+3\varepsilon - 2n\varepsilon - 2\varepsilon^2} & -\frac{c\varepsilon - b\varepsilon + b\varepsilon^2 - bn + cn + bn\varepsilon - 2cn\varepsilon}{n+3\varepsilon - 2n\varepsilon - 2\varepsilon^2} \\
GTFT & \frac{n}{n+\varepsilon} (b - c) & \frac{n}{n+\varepsilon} (b - c)
\end{bmatrix}
\]  

(5.12)

For \( n > \varepsilon \), straightforward manipulations show that:

\[
\frac{b\varepsilon - c\varepsilon + c\varepsilon^2 + bn - cn - 2bn\varepsilon + cn\varepsilon}{n+3\varepsilon - 2n\varepsilon - 2\varepsilon^2} > \frac{1}{2}b - \frac{1}{2}c \Leftrightarrow b > b^{**} = \frac{c}{1 - n - \varepsilon}
\]

\[
\frac{n}{n+\varepsilon} (b - c) > -\frac{c\varepsilon - b\varepsilon + b\varepsilon^2 - bn + cn + bn\varepsilon - 2cn\varepsilon}{n+3\varepsilon - 2n\varepsilon - 2\varepsilon^2} \Leftrightarrow b > b^* = \frac{c}{1 - 2\varepsilon}
\]

So, for \( b > b^{**} \) the game is dominance solvable with \( (GTFT, GTFT) \) as the unique Nash equilibrium\(^{11}\). Consequently, in the \( \beta \to \infty \) limit, the fraction of GTFT players converges to 1 as there is no selection pressure against a more "generous" reciprocative type as long as the rest of the population proceeds to a mimicking behavior. GTFT is a generalization of TFT and a population of TFT provides a stepstone for "nicer" behaviour to flourish. The lack of strongly defective players creates a payoff advantage for GTFT, inducing more switching away from TFT, with population eventually ending up in an all GTFT state (Fig. 5.8a).

For \( b \in (b^*, b^{**}) \) the best responses are \( TFT \in BR\{TFT\}, GTFT \in BR\{GTFT\} \) and we are in the case of a Coordination game with Logit Dynamics displaying co-existence of behavioral rules, but with GTFT largely outnumbering TFT (Fig. 5.8b).

The transition in the space of \( 2 \times 2 \) games is illustrated by the bifurcation diagram

\(^{11}\)similarly, if \( b \in (b^*, b^{**}) \) then we have a coordination game while for \( b < b^* \) the game is again dominance solvable, with \( (TFT, TFT) \) unique pure-strategy Nash equilibrium.
in Fig. 5.8f. For a tiny region of the benefit of cooperation parameter \( b \in (1, b^*)^{12} \) the game is dominance solvable with \( GTFT \) strictly dominated by \( TFT \) (see a blow-up of the diagram around \( b = 1 \) in Panel (e) where the equilibrium fraction of GTFT is close to zero before the 2-cycle emerge). Then, for the interval \( b \in (b^*, b^{**}) \) we have a coordination game with two equilibria while for \( b > b^{**} \) a dominance solvable game with \( TFT \) strictly dominated by \( GTFT \).

Finally, for a given game - \( b = 2 \), i.e. a Coordination game- letting the intensity of choice vary also leads to a transition from a unique polymorphic steady state (low values of \( \beta \)) through a 2-cycle for intermediate \( \beta \) and back to a unique monomorphic steady state when ‘close’ to best-reply dynamics.(Fig. 5.8d). A similar transition to a two-cycle is observed in Fig. 5.8e if we let the degree of generosity parameter \( n \) vary.

\(^{12}b < 1\) is not a PD game any longer, as \( c = 1. \)
Figure 5.8: Reciprocators(TFT) vs. Generous reciprocators(GTFT). This ecology gives rise to a dominance solvable game for relatively large benefits of cooperation (Panel (a)) and to a Coordination game for intermediate values (Panel (b)). For a fixed Coordination game, Panel (d) shows the transition from a unique polymorphic steady state to a 2-cycle and back to a monomorphic steady state. In Panels (c)-(f) the transition in the space of $2 \times 2$ games is illustrated for the $\beta$ large limit. Game parameters: $b = 2, c = 1, \varepsilon = 0.01, n = 0.3, \beta = 700$. 

(a) $b = 2, \beta = 700$. Time series

(c) $b = 2$. Bifurcation diagram (GTFT,$n$)

(e) $\beta = 1000$. Game transition DS-HD

(b) $b = 1.5, \beta = 700$. Time series

(d) $b = 2$. Bifurcation diagram (GTFT,$\beta$)

(f) $\beta = 700$. Game transition DS-HD-DS
5.3.9 TFT vs. WSLS

For this ecology of strategies the payoff matrix (5.3) boils down to:

\[
\begin{bmatrix}
TFT & WSLS \\
TFT & \frac{1}{2}b - \frac{1}{2}c & \frac{1}{2}b - \frac{1}{2}c \\
WSLS & \frac{1}{2}b - \frac{1}{2}c & (b - c) (1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon)
\end{bmatrix}
\]

(5.13)

For small implementation errors \(\varepsilon\), WSLS weakly dominates TFT and, in the large \(\beta\) limit, Logit Dynamics leads to elimination of weakly dominated strategies (Fig. 5.9a). Intuition draws on Pavlov essentially being a performance-dependent strategy (i.e. change status-quo when unsatisfied with the past period payoff) relative to TFT’s merely mimicking opponent past move.

For smaller \(\beta\) the payoff differential between the two strategies is small and the system converges to a steady state with co-existing rules (Fig. 5.9b).

![Time series plots](image)

(a) \(\beta = 100\). Time series  
(b) \(\beta = 1\). Time series

Figure 5.9: Reciprocators (TFT) vs. Pavlov (WSLS). Panel (a): large values of the intensity of choice, quick convergence to monomorphic population consisting only of Pavlovian players, irrespective of the cooperation benefits size \(b\). Panel (b) shows convergence to a polymorphic state for low \(\beta\). Game parameters: \(b = 2, c = 1, \varepsilon = 0.01, n = 0.3, \beta = 100\).
5.3.10 WSLS vs. AllC

The payoff matrix for this ecology reads:

\[
\begin{bmatrix}
WSLS & AllC \\
WSLS & (1 - \varepsilon)(1 - 2\varepsilon + 4\varepsilon^2)(b - c) & b - \frac{1}{2}c - b\varepsilon \\
AllC & b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2 & (1 - \varepsilon)(b - c)
\end{bmatrix}
\] (5.14)

If \( b > b^{*} \) the game (5.14) is of Hawk-Dove type with two asymmetric pure strategy equilibria \((\text{AllC}, \text{WSLS})\) and \((\text{WSLS}, \text{AllC})\) and one , interior equilibrium, in mixed strategies. One could notice that, in the \( \varepsilon \to 0 \) limit, Pavlov and AllC perform equally well against their own type but Pavlov does relatively better in a mixed matching \((b - \frac{1}{2}c > b - c)\). However for strictly positive errors, AllC does better against its own type than Pavlov \(((1 - 2\varepsilon + 4\varepsilon^2) < 1)\). This gives a payoff advantage for AllC which ensures that it does not go extinct but survives in small fractions subjected to perpetual oscillations (Fig. 5.10a). For a given game matrix, the two-cycle is created, as usual, via a period-doubling bifurcation when, for instance, the intensity of choice \( \beta \) is increased (Fig. 5.10d). Panel (c) in Fig. 5.10 displays the transition, in the \( 2 \times 2 \) games space, from a dominance solvable (low \( b \)) to a Hawk-Dove game (large \( b \)). The logit dynamics attracting set shifts from a unique monomorphic steady state with Pavlov players only to a 2-cycle alternating between WSLS and AllC.

\(^{13}b^{*} = 2c \frac{1 - \varepsilon}{1 - 2\varepsilon} \approx 2c\). Otherwise, the game is dominance solvable with WSLS surviving elimination of strictly dominated AllC.
Figure 5.10: Pavlov (WSLS) vs. Unconditional cooperators (AllC). Oscillatory co-existence of strategies with Pavlovian behaviour outnumbering unconditional cooperators for small intensity of choice (Panel (a)) and continuous swings between monomorphic states when $\beta$ approaches the "best response" limit (Panel (b)). Panel (c) illustrates the bifurcation in the space of games, as the benefit of cooperation varies, from a unique steady state (equilibrium) to a 2-cycle (2 asymmetric equilibria). Panel (d): as the intensity of choice increases the long-run outcome changes from a polymorphic steady state to a 2-cycle. Remaining game parameters: $b = 2, c = 1, \varepsilon = 0.01, n = 0.3, \beta = 15$. 

(b) $\beta = 200$. Time series

(d) $b = 2$. Bifurcation diagram (WSLS, $\beta$)

(c) $\beta = 500$. Bifurcation diagram (WSLS, $b$)
5.3.11 Summary

Table 5.1 summarizes the normal form games resulting from 10 different the strategic pairwise interaction of the five repeated PD strategies together with the long-run attractors of the corresponding evolutionary game under the perturbed best-reply dynamics. In the limiting $\beta \to \infty$ case, these long-run outcomes coincide with the attracting sets of the unperturbed best-reply dynamics on the respective $2 \times 2$ game.

<table>
<thead>
<tr>
<th>No.</th>
<th>2x2</th>
<th>Game</th>
<th>Bifurcation</th>
<th>Attractors</th>
<th>Path-Dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AllD-TFT</td>
<td>WDS,C</td>
<td>SN</td>
<td>multiple SS</td>
<td>yes, $\varepsilon = 0.05$</td>
</tr>
<tr>
<td>2</td>
<td>TFT-AllC</td>
<td>HD</td>
<td>PD</td>
<td>2-cycle</td>
<td>no</td>
</tr>
<tr>
<td>3</td>
<td>AllD-AllC</td>
<td>DS</td>
<td>none</td>
<td>unique SS</td>
<td>no</td>
</tr>
<tr>
<td>4</td>
<td>AllD-GTFT</td>
<td>C</td>
<td>SN</td>
<td>multiple SS</td>
<td>yes, $\varepsilon = 0.01$</td>
</tr>
<tr>
<td>5</td>
<td>AllD-WSLS</td>
<td>DS,C</td>
<td>SN</td>
<td>multiple SS</td>
<td>yes, $b = 4$</td>
</tr>
<tr>
<td>6</td>
<td>GTFT-AllC</td>
<td>HD</td>
<td>PD</td>
<td>2-cycle</td>
<td>yes</td>
</tr>
<tr>
<td>7</td>
<td>GTFT-WSLS</td>
<td>C</td>
<td>SN</td>
<td>multiple SS</td>
<td>yes</td>
</tr>
<tr>
<td>8</td>
<td>TFT-GTFT</td>
<td>DS, C</td>
<td>PD</td>
<td>2-cycle</td>
<td>no</td>
</tr>
<tr>
<td>9</td>
<td>TFT-WSLS</td>
<td>WDS</td>
<td>none</td>
<td>unique SS</td>
<td>no</td>
</tr>
<tr>
<td>10</td>
<td>WSLS-AllC</td>
<td>DS,HD</td>
<td>PD</td>
<td>2-cycle</td>
<td>no</td>
</tr>
</tbody>
</table>

Table 5.1: Summary of the type of game and dynamical behaviour for all 10 pairwise interaction of iterated Prisoner’s Dilemma strategies under logit dynamics.

We have encountered the following typical pairwise interactions between iterated Prisoner’s Dilemma strategies together with transition from one game form to another as the payoff matrix parameters $(b, c, \varepsilon, n)$ change:

- Coordination game (C): multiplicity of monomorphic (only one type of behaviour surviving) steady states with strong path-dependence

- Hawk-Dove game (HD): asymmetric equilibria; 2-cycle with continuous popu-
lation switching from one monomorphic state to the other

- Prisoner's Dilemma (PD): uniquely selected long run symmetric, Pareto-inferior, equilibrium (i.e. monomorphic population state)

- Dominance Solvable game (DS), non-PD: unique, monomorphic long-run steady state

- Games with a weakly dominated strategy (WDS): unique, monomorphic long-run steady state

The co-existing steady states and 2-cycles emerging under logit dynamics are usually created via fold/ saddle-node (SN) and period-doubling (PD) bifurcations, respectively. Although the qualitative changes in the space of $2 \times 2$ games do not make the object of the present chapter, some interesting game transitions were also revealed and it would be worth studying them systematically, as future research.
5.4 $3 \times 3$ Ecologies

In this section we will discuss some typical examples out of the $\binom{5}{3} = 10$ sub-ecologies of three iterated PD strategies, focusing the analysis on the resulting $3 \times 3$ normal form games and on the attracting sets and long-run behavior of the logit evolutionary dynamics.

5.4.1 AllD-TFT-AllC

This ecology has been thoroughly investigated in Sigmund and Brandt (2006) under the Replicator Dynamics. The reduced $3 \times 3$ payoff matrix reads:

\[
\begin{bmatrix}
    \text{AllD} & \text{TFT} & \text{AllC} \\
    \text{AllD} & \varepsilon (b - c) & -\varepsilon (c - 2b + 2\varepsilon) & b - b\varepsilon - c\varepsilon \\
    \text{TFT} & \varepsilon (b - 2c + 2c\varepsilon) & \frac{1}{2} b - \frac{1}{2} c & b - c - b\varepsilon + 2c\varepsilon - 2c\varepsilon^2 \\
    \text{AllC} & b\varepsilon - c + c\varepsilon & b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2 & (1 - \varepsilon) (b - c)
\end{bmatrix}
\]

After normalization (all elements on the main diagonal set equal to zero by, for instance, deducting the diagonal entry from each column for player I/or from each row for player II) we get:

\[
\begin{bmatrix}
    \text{AllD} & \text{TFT} & \text{AllC} \\
    \text{AllD} & 0 & \frac{1}{2} (1 - 2\varepsilon) (c - b + 2b\varepsilon) & c (1 - 2\varepsilon) \\
    \text{TFT} & -c\varepsilon (1 - 2\varepsilon) & 0 & c\varepsilon (1 - 2\varepsilon) \\
    \text{AllC} & c (2\varepsilon - 1) & \frac{1}{2} (1 - 2\varepsilon) (b - c + 2b\varepsilon) & 0
\end{bmatrix}
\]

As $\frac{1}{2} (1 - 2\varepsilon) (c - b + 2b\varepsilon) < 0$ and $c (1 - 2\varepsilon) > 0$ this matrix has, in the $\varepsilon \to 0$ limit, a generalized Rock-Paper-Scissors structure\textsuperscript{14}. Namely, AllD ties with its own

\textsuperscript{14}Sigmund and Brandt (2006) also discover, when mistakes probability $\varepsilon \to 0$, a RSP pattern under Replicator Dynamics, for a similar ecology of IPD game rules, with discounting.
type, loses to \(TFT\) and wins at \(AllC\).

Putting together insights from subsections (5.3.1), (5.3.2) and (5.3.3) we notice that, in a stochastic strategy framework, for \(\varepsilon \to 0\), this ecology resembles very much a Rock-Paper-Scissors pattern of behavior: each of the three repeated strategies provide the environment in which another one can emerge and take over in the following order: a population of unconditional defectors \(AllD\) is prone to invasion and supersession by small clusters of reciprocators \(TFT\) which, upon reaching a critical threshold, provide the stepstone for the emergence of unconditional cooperators \(AllC\). These \(AllC\) players grow in the population up to the limit when conditions are created for unconditional defectors \(AllD\) to thrive, strike back in the population and, thus, complete the full cycle (see Fig. 5.11e). Any small number of \(AllD\) players is able to invade a population of \(AllC\) given that \(AllD\) strongly dominates \(AllC\) (from subsection (5.3.3)) while, in fact, \(TFT\) can never invade a population of unconditional cooperators (subsection (5.3.2)). Fig. 5.11c-d depict projections of one such invariant curve onto subspaces of two strategies, together with the Hopf bifurcation route to periodic behavior (see Panels (a),(b), (f) for bifurcation diagrams of \(AllD\) and \(TFT\) fractions varying the benefit of cooperation \(b\), mistake probability \(\varepsilon\) and intensity of choice \(\beta\) parameters, respectively). The limit cycles born at finite values of \(\beta\) turn into a 3-cycle in the vanishing errors and "best-response" limits \(\varepsilon \to 0, \beta \to \infty\), according to the best-reply structure of the underlying RSP game.
Figure 5.11: AllD vs. TFT vs. AllC. Bifurcation diagrams with respect to the benefit of cooperation parameter $b$ (Panel (a)), probability of mistake (Panel (b)) and intensity of choice (Panel (f)). Panels (c)-(d): phase plots in AllD-TFT and TFT-AllC subspaces, displaying Rock-Paper-Scissors type of long-run behaviour. Last, Panel (e) shows fractions evolution for large $\beta$. Baseline game parameters: $b = 2, c = 1, \varepsilon = 0.01, n = 0.3, \beta = 10$. 
5.4.2 AllD-GTFT-WSLS

The 3x3 payoff matrix reads:

\[
\begin{bmatrix}
    \text{AllD} & \text{GTFT} & \text{WSLS} \\
    \text{AllD} & \varepsilon (b - c) & m_{13} & \frac{1}{2}b - c\varepsilon \\
    \text{GTFT} & m_{31} & \frac{n}{n+\varepsilon} (b - c) & m_{34} \\
    \text{WSLS} & b\varepsilon - \frac{1}{2}c & m_{43} & (b - c) (1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon)
\end{bmatrix}
\]  

(5.16)

With the help of the pairwise payoff comparisons\textsuperscript{15} performed in subsections (5.3.4), (5.3.5) and (5.3.7), given \( n > \varepsilon \) (by assumption), we can easily derive the best-response correspondences in game form (5.16) as follows: \{\text{AllD}\} = BR\{\text{AllD}\}, \{\text{GTFT}\} = BR\{\text{GTFT}\} and \{\text{WSLS}\} = BR\{\text{WSLS}\}. Thus, the game is a 3 x 3 coordination game, and, in the best-reply limit of our logit dynamic(i.e. \( \beta \to \infty \)) each of these three pure strategy symmetric equilibria could be obtained as long-run outcome if the population starts out with the appropriate distribution of fractions. See Fig.5.12f for a set of initial conditions from which the (WSLS, WSLS) is selected.

For \( \beta \) finite, the situation becomes more intricate as the Logit dynamics appear to break the typical coordination game best-response structure described above and generate Rock-Paper-Scissors patterns of cyclical oscillations irrespective of the initial mixture of the population (Fig.5.12a-d).

As the bifurcation diagram with respect to \( \beta \) (Fig.5.12e) suggests, a unique, fully mixed steady state for low values, destabilizes, for moderate values of the intensity of choice, via a Neimark-Sacker (NS) bifurcation, and RSP-like stable, limit cycles appear. For high values of \( \beta \) a sequence of saddle-node (SN) bifurcations occurs, conducing to three\textsuperscript{16} co-existing stable steady states, in the \( \beta \to \infty \) limit. This bifurcation scenario is confirmed in Fig. 5.13 where we continue a barycentrical fixed

\textsuperscript{15}i.e. the best-response functions for the corresponing 2 x 2 games.

\textsuperscript{16}Panel (e) displays, for large \( \beta \) only two values of Pavlov fractions steady states, the top one with 100% share of WSLS and the bottom one with 0%. The third steady state supersedes the bottom curve, with zero WSLS fraction, too.
point \((1/3, 1/3, 1/3)\) in the benefit of cooperation/error in strategy execution space (top panels) and detect the conjectured Neimark-Sacker (NS)-Period-Doubling (PD)-Saddle-Node (LP) sequence of bifurcations along the equilibrium curve. The middle and bottom panels of Fig. 5.13 ‘continue’ each detected codimension I singularity with respect to another game or behavioral parameter and displays the resulting bifurcation curves. When a NS/PD/LP curve of codimension I bifurcations is crossed from below a limit cycle/2-cycle/co-existing stable and unstable steady states emerge, respectively. Also, it is worthwhile pointing out the rich selection of codimension II bifurcations occurring along the curves of codimension I bifurcations in this 3x3 ecology of repeated Prisoner’s Dilemma rules under the Logit Dynamics.

The intuition for the emergence, within a Coordination game, of the Rock-Paper-Scissors cycles "GTFT beats AllD, W SLS beats GTFT and AllD beats W SLS" is not immediately clear. We know that in the \(\beta \to \infty\) limit the interaction is of Coordination game type with the three co-existing stable steady states created via a sequence of saddle-node bifurcations as shown before. Still, for finite values of the intensity of choice \(\beta\) (Fig.5.12e) the typical Coordination game (each strategy \(E_i\) is a best-reply to itself) behaves dynamically as if a RSP game (each strategy \(E_i\) is a best reply to its successor \(E_{i+1}\)). Nevertheless, even for finite \(\beta\) the Coordination game dynamic behavior can be recovered if the benefits accrued to cooperation \(b\) are high enough (Fig.5.12a).
Figure 5.12: AllD vs. GTFT vs. WSLS. Bifurcation diagrams with respect to the benefit of cooperation parameter $b$ (Panel (a)) and probability of mistakes (Panel (b)). Phase plots in GTFT-WSLS space (Panel (c)) and time series (Panel (d)), displaying Rock-Paper-Scissors type of long-run behaviour for moderate $\beta$. Bifurcation diagram with respect to $\beta$ and time series (Panel (f)) showing convergence, in the $\beta$ large limit to the all Pavlov steady state (for an appropriate initial population mixture). Baseline game parameters: $b = 2, c = 1, \varepsilon = 0.01, n = 0.1, \beta = 15$. 

Baseline game parameters:

(a) $\beta = 15$. Bifurcation diagram, $b$
(b) $\beta = 15$. Bifurcation diagram, $\varepsilon$
(c) $b = 2.16$. Phase plot
(d) $b = 2.16$. Time series
(e) $b = 2.2$. Bifurcation diagram
(f) $\beta = 100$. Initial fractions (0.2,0.3,0.5)
Figure 5.13: AllD vs. GTFT vs. WSLS. Top panels: continuation of an equilibrium in the benefit of cooperation and error probability space. Middle and bottom panels: continuation of the Period-Doubling (PD), Saddle-Node (LP) and Neimark-Sacker (NS) singularities detected in the top two panels with respect to a second parameter: intensity of choice $\beta$, degree of generosity $n$, benefit of cooperation $b$ and error in implementation $\epsilon$. Baseline game parameters: $b = 2, c = 1, \epsilon = 0.01, n = 0.1, \beta = 15$.  

138
5.4.3 AllD-GTFT-AllC

The restriction of the full payoff matrix (5.3) to this $3 \times 3$ sub-ecology, can be arranged, after the normalization of diagonal payoffs into the following form:

$$
\begin{bmatrix}
    \text{AllD} & \text{GTFT} & \text{AllC} \\
    \text{AllD} & 0 & -\frac{1}{n+\varepsilon} (n - \varepsilon^2 - n\varepsilon) (b - c - bn - b\varepsilon) & c(1 - 2\varepsilon) \\
    \text{GTFT} & -c(n - \varepsilon^2 - n\varepsilon) & 0 & c\varepsilon (1 - n - \varepsilon) \\
    \text{AllC} & -c(1 - 2\varepsilon) & \frac{\varepsilon}{n+\varepsilon} (1 - n - \varepsilon) (b - c - bn - b\varepsilon) & 0
\end{bmatrix}
$$

(5.17)

It is easy to see that, for $\varepsilon, n \to 0$ this normalized payoff matrix has a generalized Rock-Paper-Scissors structure, i.e. AllD outcompetes AllC, AllC outcompetes GTFT and GTFT outcompetes AllD. In this case the logit dynamics has a limit cycle as an attracting set for moderate $\beta$ and a three cycle with perpetual switching between the three monomorphic steady states in the $\beta$ large limit. For strictly positive perturbations $\varepsilon, n$ the game is not RSP per se but, nevertheless, the Logit Dynamics generates, for low to intermediate values of the intensity of choice, periodic behavior resembling the described RSP cycles. Reasoning for the emergence of such limit cycles in Fig. 5.14c-d goes along the best-reply correspondences argument already sketched in subsection (5.4.1). Unlike the AllD – TFT – AllC ecology, the shape of the curve, as well as the oscillations bounds, are distorted, with the original ones recovered in the limit $n \to \varepsilon$ when GTFT approaches the TFT strategy. However, in the $\beta \to \infty$ limit system converges to a monomorphic AllD state (Fig. 5.14e-f). The presence of undiscriminating cooperators leads to the extinction of the generous reciprocators.
Figure 5.14: AllD vs. GTFT vs. AllC. Bifurcation diagrams with respect to the benefit of cooperation parameter $b$ (Panel (a)) and degree of generosity $n$ (Panel (b)). Phase plots in AllC-GTFT space together with Rock-Paper-Scissors time series are reported in Panels (c) and (d), respectively. Panels (e),(f) display the long run behaviour in the $\beta$ large limit.

Baseline game parameters: $b = 2, c = 1, \varepsilon = 0.01, n = 0.1, \beta = 15,$
5.4.4 AllD-TFT-WSLS

The payoff matrix for this sub-ecology takes the following form:

\[
\begin{bmatrix}
\text{AllD} & \text{TFT} & \text{WSLS} \\
\text{AllD} & 0 & (\frac{1}{2} - \varepsilon) (c - b + 2b\varepsilon) & (\varepsilon - \frac{1}{2}) (b - 2c - 4b\varepsilon + 4c\varepsilon + 4b\varepsilon^2 - 4c\varepsilon^2) \\
\text{TFT} & c\varepsilon (2\varepsilon - 1) & 0 & \frac{1}{2} (b - c) (2\varepsilon - 1)^3 \\
\text{WSLS} & c (\varepsilon - \frac{1}{2}) & 0 & 0 \\
\end{bmatrix}
\]

which, in the limit \(\varepsilon \to 0\), simplifies to:

\[
\begin{bmatrix}
\text{AllD} & \text{TFT} & \text{WSLS} \\
0 & -\frac{1}{2}(b - c) & c - \frac{1}{2}b \\
0 & 0 & -\frac{1}{2}(b - c) \\
-\frac{1}{2}c & 0 & 0 \\
\end{bmatrix}
\]

For relatively small benefits of cooperation \((c < b \leq 2c)\) we see that \text{AllD} loses to \text{TFT} \((\frac{1}{2} - \varepsilon) (c - b + 2b\varepsilon) < 0\). Unconditional defectors \text{AllD} win against \text{WSLS} \((\varepsilon - \frac{1}{2}) (b - 2c - 4b\varepsilon + 4c\varepsilon + 4b\varepsilon^2 - 4c\varepsilon^2) > 0\) and tie against own-type. Thus, this matrix has a degenerate generalized Rock-Paper-Scissors structure and limit cycles are born via a Hopf bifurcation, for low\(^\text{17}\) to moderate values of the intensity of choice parameter \(\beta\) (Fig. 5.15a-d) with a 3-cycle arising in the \(\beta \to \infty\) limit. For strictly positive mistake probability \(\varepsilon\) there are still RPS cycles born, but only for moderate values of the intensity of choice, while, for large intensity of choice, the population ends up in an \text{AllD} monomorphism (Panel(e)). The intuition for these oscillations could be extracted from results derived in subsections (5.3.1), (5.3.5) and (5.3.9): for small errors of implementation \text{TFT} weakly dominates \text{AllD}, in turn, is it weakly dominated by \text{WSLS} and, finally, \text{AllD} easily invades a Pavlovian population when benefits accrued to cooperation are small (see Fig. 5.5a). For strictly positive \(\varepsilon\) and

\(^{17}\text{If} \beta = 0, \text{the repeated strategy choice is random and the unique steady state is the fully mixed population state.}\)
\( \beta \to \infty \) the population converges to an \textit{All}D monomorphic state.

If the relative benefits are large \((b > 2c)\), the RSP structure disappears for \( \varepsilon \to 0 \), and, for \( \beta \) large the population converges, irrespective of the initial mixture, to a Pavlov-only state (Panel (f)). Notice that, even if the underlying game itself is no longer RSP, there are still oscillations for moderate \( \beta \) due to the evolutionary dynamics.
Figure 5.15: AllD vs. TFT vs. WSLS. Bifurcation diagrams with respect to the benefit of cooperation parameter $b$ (Panel (a)) and probability of mistakes (Panel (b)). Phase plots in AllD-WSLS (Panel (c)) and time series (Panel (d)) displaying Rock-Paper-Scissors type of long-run behaviour. Panels (e)-(f) show long-run behavior and convergence, for large $\beta$, to AllD (Pavlov) monomorphism for $b < 2c$ ($b > 2c$, respectively). Baseline game parameters: $b = 2, c = 1, \varepsilon = 0.01, n = 0.2, \beta = 15$, $\delta = 0.01$. 

143
5.4.5 AllD-TFT-GTFT

The reduced $3 \times 3$ payoff matrix for an ecology of defectors and (generous) reciprocators reads:

\[
\begin{bmatrix}
\text{AllD} & \text{TFT} & \text{GTFT} \\
\varepsilon (b - c) & -\varepsilon (c - 2b + 2\varepsilon) & b\varepsilon - c\varepsilon - b\varepsilon^2 + bn - bn\varepsilon \\
\varepsilon (b - 2c + 2c\varepsilon) & \frac{1}{2}b - \frac{1}{2}c & \frac{-c\varepsilon - 2b\varepsilon^2 - 2bn - 2n\varepsilon - 2c\varepsilon}{n + 3\varepsilon - 2n\varepsilon - 2\varepsilon^2} \\
b\varepsilon - c\varepsilon + c\varepsilon^2 - cn + cn\varepsilon & \frac{b\varepsilon - c\varepsilon + c\varepsilon^2 + bn - cn - 2bn\varepsilon + cn\varepsilon}{n + 3\varepsilon - 2n\varepsilon - 2\varepsilon^2} & \frac{n}{n + \varepsilon} (b - c)
\end{bmatrix}
\]

(5.19)

Using the inequalities from the relevant $2 \times 2$ sub-ecologies (subsections (5.3.1), (5.3.4) and (5.3.8)) we can show that, for $b > \frac{c}{1-n}$, there are only two pure strategy equilibria (AllD, AllD) and (GTFT, GTFT). By adding the GTFT strategy to the coordination game AllD vs. TFT in (5.3.1), the steady state (TFT, TFT) is lost. However, TFT is not a dominated strategy and it plays an ‘intermediate layer’ role in ensuring transition from AllD to GTFT. Moreover, for small $\varepsilon$, TFT breaks the path-dependence observed in the AllD vs. GTFT ecology (see subsection (5.3.4)) and, in the $\beta \to \infty$ limit, (GTFT, GTFT) is the unique steady state (Fig. 5.16b,f).

The intuition relies on the fact that there are only two possible transitions leading to an GTFT—only population: from TFT to GTFT (‘immediate’ in the $\beta$ large case as TFT is strictly dominated strategy) and from AllD to GTFT (in two stages): in the first stage TFT (unlike GTFT) invades and supersedes AllD for any initial fraction of defectors while in the second stage GTFT supersedes TFT according to the strict dominance criterion. Interestingly, for relatively small benefit of cooperation\textsuperscript{18}, oscillatory co-existence of all three strategies (Fig. 5.16c-d) is observed for moderate $\beta$, while for $\beta \to \infty$ population converges to an AllD state.

\textsuperscript{18} Actually, the normalized game matrix in the $\varepsilon \to 0$ limit, is

\[
\begin{bmatrix}
0 & \frac{1}{2}c - \frac{1}{2}b & c - b + bn \\
0 & 0 & 0 \\
-cn & \frac{1}{2}b - \frac{1}{2}c & 0
\end{bmatrix}
\]

a degenerate Rock-Paper-Scissors for $c < b < \frac{c}{1-n}$.
Figure 5.16: AllD vs. TFT vs. GTFT. Panels (a)-(d): Evolution of strategies fractions for small and large benefit of cooperation coupled with low and high values of the sensitivity to payoff differences, respectively. Remaining (game) parameters set to: $c = 1$, $\varepsilon = 0.01$, $n = 0.3$. 

(a) $b = 3$, $\beta = 25$. Time series

(b) $b = 3$, $\beta = 300$. Time series

(c) $b = 1.2$, $\beta = 29.8$. Limit cycle

(d) $b = 1.2$, $\beta = 300$. 5-cycle

(e) $b < c/(1 - n)$. Bifurcation diagram

(f) $b > c/(1 - n)$. Bifurcation diagram
5.4.6 AllD-WSLS-AllC

The reduced 3 × 3 payoff matrix for an ecology of defective, stimulus-response and cooperative players takes the following form:

\[
\begin{bmatrix}
AllD & W S L S & AllC \\
AllD & \varepsilon(b-c) & \frac{1}{2}b-c\varepsilon & b-b\varepsilon-c\varepsilon \\
W S L S & b\varepsilon-\frac{1}{2}c & (b-c)(1-4\varepsilon^3+6\varepsilon^2-3\varepsilon) & b-\frac{1}{2}c-b\varepsilon \\
AllC & b\varepsilon-c+c\varepsilon & b-c-2b\varepsilon+c\varepsilon+2b\varepsilon^2 & (1-\varepsilon)(b-c)
\end{bmatrix}
\] (5.20)

We notice first, that AllC is strictly dominated by W S L S provided that \( b < 2c \frac{1-\varepsilon}{1-2\varepsilon} \), the necessary and sufficient condition for Pavlov to be a best-reply to itself. In this case the equilibria structure is given by the equilibria of the AllD vs. W S L S game - (AllD, AllD) and (W S L S, W S L S) - in subsection (5.3.5). But, in the limit \( \varepsilon \to 0 \), the ecology AllD - W S L S was dominance solvable for \( b < 2\varepsilon \) with W S L S the strictly dominated strategy. Combining the two restrictions, we obtain (AllD, AllD) as the only remaining Nash equilibrium in our 3x3 ecology for small benefits of cooperation \( b < 2c \).

If \( b > 2c \frac{1-\varepsilon}{1-2\varepsilon} \), then AllC is a best-reply to W S L S, but W S L S is not a best reply to AllC (as \( b-b\varepsilon-c\varepsilon > b-\frac{1}{2}c-b\varepsilon \)) so both the Coordination and the Hawk-Dove patterns of the AllD - W S L S and W S L S - AllC ecologies (see subsection (5.3.10) are broken. Again we are only left with (AllD, AllD) as the unique pure strategy equilibrium.(Fig. 5.17a). We observe the same phenomenon as with the AllD - GFTF - AllC ecology, i.e. the presence of undiscriminating players AllC leads to the extinction of the discriminating type Pavlov.

Interestingly, if we relax the \( \beta \to \infty \) assumption, the logit dynamics easily generates periodic (Fig. 5.17b-c) or chaotic behavior with irregular switching between the three strategies. Fig. (5.17d) shows a plot of the largest Lyapunov exponent λ, with \( \lambda > 0 \) implying chaos. In this case complicated dynamics arises through a period-doubling route to chaos.
Figure 5.17: AllD vs. WSLS vs. AllC. Time series evolution for small and high values of $\beta$ (Panels (a)-(b). Bifurcation diagram with respect to the benefit of cooperation $b$ (intensity of choice $\beta$) and evidence for chaotic behavior in Panels (c)-(d) and (e)-(f), respectively. Game parameters: $b = 4, c = 1, \varepsilon = 0.01, n = 0.3$. 

\( a \) $b > 2c, \beta = 300$. Time series

\( b \) $b > 2c, \beta = 30$. Time series

\( c \) $\beta = 30$. Bifurcation diagram

\( d \) $\beta = 30$. Largest Lyapunov exponent

\( e \) $b = 4.4$. Bifurcation diagram

\( f \) $b = 4.4$. Largest Lyapunov exponent
5.4.7 TFT-WSLS-AllC

The 3 × 3 payoff matrix for an ecology of reciprocators, Pavlovians and cooperators is given by:

\[
\begin{bmatrix}
TFT & WSLS & AllC \\
\frac{1}{2}b - \frac{1}{2}c & \frac{1}{2}b - \frac{1}{2}c & b - c - b\varepsilon + 2c\varepsilon - 2c\varepsilon^2 \\
\frac{1}{2}b - \frac{1}{2}c & (b - c) (1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon) & b - \frac{1}{2}c - b\varepsilon \\
b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2 & b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2 & (1 - \varepsilon)(b - c)
\end{bmatrix}
\]

(5.21)

By comparing entries in this payoff matrix, we notice the best-reply correspondences: \{AllC\} = BR\{TFT\}, \{WSLS\} = BR\{AllC\}. Also, \{AllC\} = BR\{WSLS\} for \(b > 2c\) and \{WSLS\} = BR\{WSLS\} if \(b < 2c\). Consequently, apart from mixed strategies equilibria, the pure strategy equilibria structure \{(WSLS, AllC), (AllC, WSLS)\} from subsection (5.3.10) is preserved for large benefit of cooperation \((b > 2c)\) while the one from subsection (5.3.9) (i.e. Coordination Game, with WSLS weakly dominating TFT), is preserved for relatively small \(b\). Moreover, TFT remains weakly dominated by WSLS even in this enlarged ecology. These two situations are illustrated in Fig. (5.18): a Hawk-Dove stable 2-cycle for large \(b\) (see Panels (a) and (b) for the unboundedly and boundedly rational cases) and a Coordination structure with a stable steady state for small \(b\) (Panels (c) and (d) for large and small \(\beta\), respectively).
Figure 5.18: TFT vs. WSLS vs. AllC. Depending on the size of the cooperation benefit two situations may arise: a Hawk-Dove pattern with complete population swings from one monomorphic state to the other (Panel (b)) or a Coordination Game with most of the initial conditions leading to a monomorphic Pavlov population (Panel (d)). Panels (a) and (c) display the boundedly rational (i.e. small intensities of choice values) versions of the two benchmarks. Remaining game parameters: $c = 1, \varepsilon = 0.01, n = 0.3$. 
5.4.8 TFT-GTFT-WSLS

The payoff matrix for this $3 \times 3$ ecology of behavioral rules reads:

$$
\begin{bmatrix}
TFT & GTFT & WSLS \\
\frac{1}{2}b - \frac{1}{2}c & -\frac{ce - bc + be^2 - bn + cn + bne - 2cne}{n + 3\epsilon - 2ne - 2\epsilon^2} & \frac{1}{2}b - \frac{1}{2}c \\
\frac{bc - ce + ce^2 + bn - cn - 2bne + cne}{n + 3\epsilon - 2ne - 2\epsilon^2} & \frac{n}{n + \epsilon} (b - c) & m_{34} \\
\frac{1}{2}b - \frac{1}{2}c & m_{43} & (b - c) (1 - 4\epsilon^3 + 6\epsilon^2 - 3\epsilon)
\end{bmatrix}
$$

(5.22)

In subsection (5.3.8) it was shown that GTFT strictly dominates TFT while subsection (5.3.9) showed that WSLS and GTFT are best responses to themselves, in the context of the corresponding 22 ecologies. Building on these results and investigating the relevant additional inequalities\(^{19}\) in the above matrix, we can derive the best-reply correspondences for the integrated $3 \times 3$ ecology as follows: $\{GTFT\} = BR\{TFT\}$, $\{GTFT\} = BR\{GTFT\}$ and $\{WSLS\} = BR\{WSLS\}$. The resulting symmetric pure strategies equilibria are: $\{GTFT, GTFT\}$ and $\{WSLS, WSLS\}$. Fig. 5.19a-b displays, in the $\beta$ large case, this path-dependence and the convergence to one of these two equilibria as we start with different population mixtures. For imperfect switching to the best-reply, i.e. for lower values of $\beta$, oscillatory dynamics may arise as well (Fig. 5.19a-b).

\(^{19}\) \(\frac{n}{n + \epsilon} (b - c) > m_{43}\) and \(\frac{n}{n + \epsilon} (b - c) > -\frac{ce - bc + be^2 - bn + cn + bne - 2cne}{n + 3\epsilon - 2ne - 2\epsilon^2}\) and \((b - c) (1 - 4\epsilon^3 + 6\epsilon^2 - 3\epsilon) > m_{34}\) from subsection (5.3.7)
Figure 5.19: TFT vs. GTFT vs. WSLS. Dependence on the population initial mixture for rational choice (Panels (a)-(b)) and oscillatory dynamics for low intensity of choice (Panel (c)). Remaining game parameters: $b = 4, c = 1, \varepsilon = 0.01, n = 0.3$. 
5.4.9 Summary

Table 5.2 summarizes the qualitative long-run behavior of logit dynamics on various $3 \times 3$ sub-ecologies of PD repeated strategies. There are two basic routes to periodic behavior: the first one is the limiting case $\varepsilon \to 0$ when five out of the ten normal form games unveil a (degenerate) generalized Rock-Paper-Scissors structure. The Logit Dynamics on this generalized RSP game gives rise to stable limit cycles created via a supercritical Neimark-Sacker (Hopf) bifurcation (see also Chapter 2) with the three monomorphisms connected via a 3-cyle in the $\beta \to \infty$ best-reply limit. In the second route to oscillating behavior, when we maintain the strictly positive error parameter $\varepsilon$, the game forms are not of RPS type, but, nevertheless, for moderate values of the intensity of choice stable oscillations created via a supercritical NS bifurcation are observed. However, for such non-RPS games, in the limiting case $\beta \to \infty$, the cyclical structure disappears with one of the strict equilibria selected as the long run outcome.

A somewhat different scenario occurs for the $AllD – GTFT – W S L S$ $3 \times 3$ co-ordination game, where a sequence of Neimark-Sacker and Fold bifurcations unfolds as the intensity of choice slides from low to high values: the long-run behavior varies from a unique polymorphic steady state for low $\beta$, through limit cycles to three co-existing steady states corresponding to the underlying game Nash equilibria in the best-reply limit.

In the transition to the $AllC$ ecologies, in the second half of the table, we have found one $3 \times 3$ example of chaotic dynamics created through a period-doubling route to chaos. Chaos appeared for finite intensity of choice $\beta$, in the $AllD – Pavlov – AllC$ ecology, whose underlying payoff matrix has a non-RSP structure even for the limiting case $\varepsilon \to 0$. The typical RSP cyclical dominance structures seem to preclude the onset of chaos.

In the last four ecologies in Table 5.2 with no defectors ($AllD$), the long-run, $\beta \to \infty$ outcomes consist, generally, of simple behavior, steady states or 2-cycle with
continuous population swings between one monomorphic state to another, with the third strategy extinct from the population.

<table>
<thead>
<tr>
<th>No.</th>
<th>3×3</th>
<th>Bifurcation</th>
<th>Attractors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AllD-TFT-AllC</td>
<td>NS</td>
<td>limit cycle</td>
</tr>
<tr>
<td>2</td>
<td>AllD-GTFT-WSLS</td>
<td>NS,LP</td>
<td>limit cycle</td>
</tr>
<tr>
<td>3</td>
<td>AllD-GTFT-AllC</td>
<td>NS</td>
<td>limit cycle</td>
</tr>
<tr>
<td>4</td>
<td>AllD-TFT-WSLS</td>
<td>NS</td>
<td>limit cycle</td>
</tr>
<tr>
<td>5</td>
<td>AllD-TFT-GTFT</td>
<td>NS</td>
<td>limit cycle</td>
</tr>
<tr>
<td>6</td>
<td>AllD-WSLS-AllC</td>
<td>PD</td>
<td>2-cycle, chaos</td>
</tr>
<tr>
<td>7</td>
<td>TFT-WSLS-AllC</td>
<td>PD</td>
<td>2-cycle</td>
</tr>
<tr>
<td>8</td>
<td>TFT-GTFT-WSLS</td>
<td>PD</td>
<td>2-cycle</td>
</tr>
<tr>
<td>9</td>
<td>GTFT-WSLS-AllC</td>
<td>PD</td>
<td>2-cycle</td>
</tr>
<tr>
<td>10</td>
<td>TFT-GTFT-AllC</td>
<td>PD</td>
<td>2-cycle</td>
</tr>
</tbody>
</table>

Table 5.2: Summary of the dynamical behaviour for all 10 3×3 interaction of iterated Prisoner’s Dilemma strategies under logit dynamics.

5.5 4×4 ecologies

Before studying the full 5×5 ecology, in this section we discuss the long-run behavior of the five 4-types interaction under the Logit Dynamics.

5.5.1 No TFT

This ecology appends the AllD – GTFT – WSLs subset (subsection (5.4.2)) with unconditional cooperators. In the β large limit, AllC destroys the Coordination game structure of AllD – GTFT – WSLs and the population converges to an AllD monomorphism, irrespective of the initial conditions (Fig. 5.20a). This is due, on the one
hand, to the Hawk-Dove nature of the $WSLS - AllC$ and $GTFT - AllC$ interactions—
which breaks the best-response correspondences in (5.4.2) $\{WSLS\} = BR\{WSLS\}$
and $GTFT = BR\{GTFT\}$ - and, on the other hand, to the unconditional defectors
$AllD$ "easiness" of invading undiscriminating cooperators $AllC$ relative to $WSLS$
and $GTFT$. Briefly, adding "cooperative" players leads to the extinction of the two
discriminating types ($GTFT$ and $WSLS$) and, consequently, the overall level of long-
run cooperation drops. However, for moderate $\beta$, Pavlov seems to be favoured by
the evolutionary dynamics\footnote{Also, from subsection (5.3.7) ($WSLS; WSLS$) equilibrium is more likely to emerge because of a
larger basin of attraction compared to ($GTFT; GTFT$); still this dependence of initial distribution
(even though basins are largely asymmetric) seems to be lost once we investigate the enlarged,
no-$TFT$ ecology.} (see Panel (e) for the evolution of an initially balanced
population) and rich dynamics of the $WSLS$ fraction unfold if either the mistake
probability ($\varepsilon$) or the intensity of choice is altered (Fig. 5.20b-c). Co-existence of
attractors - one chaotic (Panel (d) plots the largest Lyapunov exponent as numerical
evidence for chaos) and an $AllD$ monomorphic steady state - is detected for inter-
mediate values of $\beta$ while the entire population becomes $AllD$ for large values of
the intensity of choice. Panels (e)-(f) show plots of fractions evolution on the two
attractors: a fully mixed population converges to the chaotic attractor (Panel (e))
while a population with no reciprocators (Panel (f)) ends up in a defectors $AllD$ only
state. This also illustrates the critical role $TFT$ players have in the emergence and
success of Pavlovian ($WSLS$) types.
Figure 5.20: AllD vs. GTFT vs. WSLS vs. AllC. Evolution towards an AllD population for $\beta$ large limit. Panels (b)-(c) plot bifurcation diagrams with respect to the error in strategy implementation $\varepsilon$ and intensity of choice $\beta$ parameters, while numerical evidence, i.e. positive Lyapunov exponent, for chaos in Panel (d). Last, evolution of fractions on the two co-existing attractors, for the boundedly rational choice, is displayed in Panels (e) and (f). Baseline game parameters: $b = 3, c = 1, \varepsilon = 0.01, n = 0.3$. 

(a) $\beta = 100$. Time series

(b) $\beta = 22$. Bifurcation diagram

(c) $b = 3$. Bifurcation diagram

(d) Largest Lyapunov exponent

(e) $\beta = 22$. Initial mix (1/4,1/4,1/4,1/4)

(f) $\beta = 22$. Initial mix (0.5,0,0.1,0.4)
5.5.2 No AllD

In the absence of hard defectors AllD, the system converges to a 2-cycle where the population displays either full swings back and forth between a Pavlovian state and an AllC state (for large $\beta$), or bounded fluctuations with Pavlov outnumbering AllC (moderate values of the intensity of choice) (Fig. 5.21a-b). The intuition for the 2-cycle as unique long-run attractors goes along the following lines: First, both GTFT and WSLS dominate TFT("strictly" or "weakly", respectively) and so TFT is not to be expected in the long run of logit dynamics. Second, each of the GTFT and WSLS generates a Hawk-Dove pattern (i.e. 2-cycle) in an ecology with AllC. Then, which of the two 2-cycles (GTFT – AllC or WSLS – AllC) will eventually emerge, depends on the evolutionary competition between GTFT, WSLS and AllC. The game matrix restricted to this ecology reads:

$$
\begin{pmatrix}
\text{GTFT} & \text{WSLS} & \text{AllC} \\
\frac{n}{n+\varepsilon} (b-c) & m_{34} & b-c - b\varepsilon + 2c\varepsilon - c\varepsilon^2 - cn\varepsilon \\
m_{43} & (b-c) (1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon) & b - \frac{1}{2}c - b\varepsilon \\
b - c - 2b\varepsilon + c\varepsilon + b\varepsilon^2 + bn\varepsilon & b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2 & (1 - \varepsilon)(b-c)
\end{pmatrix}
$$

The following inequalities are helpful for constructing the best-response correspondences:

$$
b - c - 2b\varepsilon + c\varepsilon + b\varepsilon^2 + bn\varepsilon > \frac{n}{n+\varepsilon} (b-c), \text{ if } b > \frac{c}{1-n-\varepsilon}
$$

$$
b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2 > (b-c)(1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon), \text{ if } b > 2c\frac{1-\varepsilon}{1-2\varepsilon}
$$

We see that for large enough benefits of cooperation, $b > 2c\frac{1-\varepsilon}{1-2\varepsilon}$, adding unconditional cooperators AllC, breaks the initial Coordination game structure of the GTFT – WSLS ecology and there are only two asymmetric pure strategy equilibria left, namely (WSLS, AllC) and (AllC, WSLS). There are never co-existing
two 2-cycles in this ecology. The WSLS – AllC 2-cycle depicted in Fig. 5.21a-b for the "boundedly rational" and the "best-reply" repeated PD strategy choices, respectively.

For \( \frac{c}{1-n-\varepsilon} < b < 2c \frac{1-\varepsilon}{1-2\varepsilon} \) we have \( AllC = BR\{GTFT\} \), \( WSLS = BR\{WSLS\} \), \( WSLS = BR\{AllC\} \) and the population converges, for \( \beta \to \infty \), to a Pavlovian WSLS monomorphism.

Last, for \( c < b < \frac{c}{1-n-\varepsilon} \), we have \( WSLS = BR\{GTFT\} \) so \( AllC \) is a strictly dominated strategy. We are back in the GTFT – WSLS ecology where Pavlov strictly dominates GTFT and, consequently, is the unique monomorphic long-run outcome of Logit Dynamics (Panels (c), (d) for high and low values of \( \beta \) respectively). We can conclude from this analysis that co-existence of two 2-cycles does not occur.
Figure 5.21: TFT vs. GTFT vs. WSLS vs. AllC. Panels (a)-(b): for large benefits of cooperation, $b = 4$, fractions converge to a (WSLS, AllC) 2-cycle for both high and low values of the intensity of choice. Panels (c)-(d): for low benefit of cooperation, $b = 1.1$, system converges to a monomorphic Pavlovian state. Remaining game parameters: $c = 1, \varepsilon = 0.01, n = 0.3$. 
5.5.3 No GTFT

Unlike the previous case, the $4 \times 4$ ecology without $GTFT^{21}$ exhibits path-dependence. Hence, for moderate values of the intensity of choice $\beta \approx 27$, the system displays co-existence of attractors. Depending on the initial population mix, the long-run evolutionary outcome can be either the cyclical competition between $AllD, TFT$ and $AllC$ with asymmetric amplitudes (time series Fig. 5.22b) as in subsection (5.4.1), or a 1-piece chaotic attractor (time series in Fig. 5.22c) with Pavlov having a dominant position, as in, for instance, subsection (5.4.6). In both cases, the other strategies remain present in the population, but in small numbers.

For high values of $\beta$ the path-dependence is lost, the only remaining attractor being a full-ranged Rock-Paper-Scissors 3-cycle (Panel (d)) in $AllD, TFT$ and $AllC$, with stimulus-response players $WSLS$ going extinct. The bifurcation diagram in Panel (a) organizes the long-run outcomes of this ecology and displays the two co-existing attractors: the chaotic one in the top and the cyclical one in the bottom curves.

---

$^{21}$Imhof et al. (2007) also discuss, in a finite population framework, an ecology of $AllD, TFT, WSLS$ and $AllC$ behaviors, under the frequency-dependent, mutation-selection Moran process. They show that $AllD$ is selected for stage game payoffs (i.e. $b/c$ ratio) below a certain threshold, while $WSLS$ wins the evolutionary competition otherwise.
Figure 5.22: AllD vs. TFT vs. WSLS vs. AllC. Panel (a) bifurcation diagram of Pavlov fractions with respect to the intensity of choice parameter; Panels (b)-(c): depending on initial distribution of fractions the long-run population state may be either ongoing oscillations among AllD, TFT, AllC(Panel (b)) or co-existence, although in a possibly chaotic fashion, of WSLS and AllC players with WSLS largely outnumbering unconditional cooperators (Panel (c)). Last, Panel (d) displays the large $\beta$ limit of the long-run dynamics of fractions. Remaining game parameters: $b = 4, c = 1, \varepsilon = 0.01, n = 0.3$. 
5.5.4 No WSLS

With the stimulus-response $WSLS$ rule put aside, the evolutionary dynamics becomes very complicated. In the $\beta$ large limit, co-existence of the four strategies in a cyclical fashion, is observed for high benefits of cooperation $b$ (Fig. 5.23a) while, for low benefits $b$, $GTFT$ is driven extinct and full-scale cyclical dynamics with the remaining three strategies emerge (Panel (b)).

For finite $\beta$, a stable limit cycle is born via a Neimark-Sacker bifurcation (Fig. 5.23a) for small $b$, while a period-doubling route to chaos emerges for high benefits accrued to cooperation (Panel (d)). Finally, in Fig. 5.23e a period-doubling route to chaos is depicted (positive Lyapunov exponent plot in Panel (f)) as the generosity parameter $n$ built into the $GTFT$ strategy is fine-tuned.
Figure 5.23: AllD vs. TFT vs. GTFT vs. AllC. Time series of the corresponding fractions, for large $\beta$, and large/small benefit of cooperation (Panel (a) and (b), respectively). Bifurcation diagrams for high and low $b$ are also shown in Panel (c)-(d). Last, Panel (e) displays a bifurcation diagram of GTFT fraction with respect to the "generosity" parameter $n$ while Panel (f) shows some numerical evidence for chaos. Remaining game parameters: $c = 1, \varepsilon = 0.01, n = 0.3$. 
5.5.5 No AllC

There are two cases to be distinguished for an AllD – TFT – GTFT – WSLS ecology. These two situations emerge from the discussion of the corresponding $3 \times 3$ sub-ecologies (see subsections (5.4.2), (5.4.4) and (5.4.5), namely low $(b \leq 2c)$ and high benefits of cooperation $(b > 2c)$.

For low $b \leq 2c$ and in the $\beta \to \infty$ limit, the system converges to a unique monomorphic GTFT state (Fig. 5.24a). However, all four behavioral rules co-exist, either in the cyclical (Panel (a) or even chaotic (Panels (b)-(d)) manner, for intermediate values of the intensity of choice.

As the responsiveness to payoffs differentials increases, quasi-periodic behavior emerges through a Neimark-Sacker bifurcation, which, if $\beta$ is pushed even further, breaks into a chaotic attractor (see Fig. 5.24b for a plot of the positive largest Lyapunov exponent). The strange attractor depicted in Panel (d) emerges via one such Neimark-Sacker, "breaking of the invariant circle" route to chaos.

When benefits of cooperation are relative large $(b > 2c)$ there is path-dependence producing the co-existence of two monomorphisms GTFT and WSLS, respectively. Panel (e) shows, for $b = 4$, the bifurcation diagram of GTFT fraction with respect to $\beta$ and we can see the two steady states created for large responsiveness to payoff differences (at the top and bottom of the panel). For instance, an initial fractions distribution biased towards Pavlov converges to a Pavlov-only state (Panel (f)). Alternatively, an even set of initial conditions favors the GTFT as the long-run, good-fated type of behavior.
Figure 5.24: AllD vs. TFT vs. GTFT vs. WSLS. Panels (a): bifurcation diagram with respect to the intensity of choice showing quasi-periodic behavior. Panels (c)-(d): time series, phase portrait and Lyapunov exponent plot as numerical evidence for chaotic time series when the intensity of choice $\beta$ increases. Finally, Panels (e)-(f) show the long-run behavior, in the best-reply limit, for large benefits of cooperation: this is, co-existence of monomorphic GTFT and Pavlov WSLS states. Baseline game parameters: $c = 1, \varepsilon = 0.01, n = 0.3$. 

(a) $b = 2$. Bifurcation ($\beta, GTFT$)  
(b) $b = 2$. Lyapunov exponent  
(c) $b = 2, \beta = 18$. Time series  
(d) $b = 2, \beta = 17.6$. Strange attractor  
(e) $b = 4$. Bifurcation diagram  
(f) $\beta = 100$. Initial mix: (0.1, 0.1, 0.1, 0.7)
5.5.6 Summary

Table 5.3 summarizes the long-run behavior of all $4 \times 4$ ecologies both the best-reply ($\beta$ large) and in the boundedly rational ($\beta$ small) choice cases. In comparison to the $3 \times 3$ ecologies reviewed in the previous section, having more strategies not only enlarges the set of possible outcomes but also makes the behavior much more complicated. In particular there is now strong path-dependence in most of the ecologies, leading to the co-existence of complicated attractors, periodic and chaotic ones. We identified two routes to complex dynamics in these $4 \times 4$ ecologies: the period-doubling and breaking of the invariant circle route to chaos. However, the path dependence is lost for some ecologies (subsections (5.5.1), (5.5.2), (5.5.3))) when $\beta$ approaches the "full rationality" limit. WSLS fares well in an AllD environment with no AllC players, but poorly, almost going extinct, in an AllC environment with AllD strategists around.

<table>
<thead>
<tr>
<th>4x4</th>
<th>Bifurcations</th>
<th>small $\beta$</th>
<th>large $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No AllD</td>
<td>PD</td>
<td>2-cycle</td>
<td>2-cycle or unique SS</td>
</tr>
<tr>
<td>No TFT</td>
<td>PD, NS</td>
<td>co-existence SS and chaos</td>
<td>unique SS (AllD)</td>
</tr>
<tr>
<td>No GTFT</td>
<td>PD, NS</td>
<td>co-existence RSP and chaos</td>
<td>unique SS (AllD)</td>
</tr>
<tr>
<td>No WSLS</td>
<td>NS</td>
<td>limit cycles, chaos</td>
<td>4-cycle</td>
</tr>
<tr>
<td>No AllC</td>
<td>NS, PD</td>
<td>limit cycles, chaos</td>
<td>multiple SS (GTFT, WSLS)</td>
</tr>
</tbody>
</table>

Table 5.3: Summary of the dynamical behaviour for all 4x4 interaction of iterated Prisoner's Dilemma strategies under logit dynamics.
5.6 5×5 Ecology

In this section, we report simulations about the behavior of Logit Dynamics on the full 5 × 5 ecology $AllD − TFT − GFT − W SLS − AllC$ that generated the game matrix (5.3). Depending on the relative size of cooperation benefits we distinguish again two situations.

First, for small benefits, $b ≤ 2c$ there is co-existence of all five strategies, albeit in a fluctuating manner, for low responsiveness to payoffs differences (Fig. 5.25a). When $β$ reaches the best-response limit the long-run outcome is an $AllD-TFT-AllC$ 3-cycle (Fig. 5.25b).

If the benefits of cooperation are large, Pavlov wins the evolutionary competition for small $β$ (Panel (c)) but goes extinct in the $β → ∞$ limit, with a 4-cycle in the remaining strategies as long run behavior (Panel (d)). Last, Panels (e), (f) display bifurcation diagrams of $AllC$ and Pavlov fractions respectively with respect to the intensity of choice. It is apparent from the $AllC$ bifurcation diagram in Panel (e) that an initially stable steady state for very low $β$, destabilizes for moderate values, with continuous swings between high and low fractions emerging for large $β$. As far as Pavlov is concerned the corresponding bifurcation diagram shows its way to extinction when players choose a best-reply (Fig. 5.25f).

The chaotic patterns in the two diagrams for intermediate values of $β$ are confirmed in Fig. 5.26a-f that displays two dimensional projections of a chaotic attractor together with the fractions evolution on the attractor for $b = 2.4805$ and $β = 15$. Finally, Fig. 5.27c-d illustrates the "breaking of an invariant circle" route to chaos: a quasiperiodic attractor for $β = 0.05$ breaks into a 6-piece quasi-periodic and then chaotic attractor (Panels (d)-(e)) at $β ≈ 10.05$ and $β = 10.2$. Eventually these pieces join together to form a 1-piece strange attractor (Fig. 5.27f) when the intensity of choice $β$ reaches the threshold $β = 10.8$. 

166
Figure 5.25: AllD vs. TFT vs. GTFT vs. WSLS vs. AllC. Panels (a)-(b) show the long-run outcome for low benefits of cooperation $b$ in the $\beta$ low and high limit, respectively. Similar time series are displayed in Panels (c)-(d) for the high benefit $b$ case. Panels (e)-(f) show, for high $b$, bifurcation diagrams of AllC and Pavlov fractions. Baseline game parameters: $c = 1, \varepsilon = 0.01, n = 0.3.$
Figure 5.26: AllD vs. TFT vs. GTFT vs. WSLS vs. AllC. Panels (a)-(d): two-dimensional projections of a strange attractor for $b = 2.4805$ and $\beta = 15$. Panels (e)-(f) evolution of fractions on the strange attractor. Remaining game parameters: $c = 1, \varepsilon = 0.01, n = 0.3$. 

(a) $b = 2.48, \beta = 15$. Strange attractor
(b) $b = 2.48, \beta = 15$. Strange attractor

c) $b = 2.48, \beta = 15$. Strange attractor
(d) $b = 2.48, \beta = 15$. Strange attractor

(e) $b = 2.48, \beta = 15$. Time series
(f) $b = 2.48, \beta = 15$. Time series
Figure 5.27: AllD vs. TFT vs. GTFT vs. WSLS vs. AllC. Panels (a)-(b): bifurcation diagram with respect to intensity of choice $\beta$ and numerical evidence of chaos (plot of the largest Lyapunov exponent). Panels (c)-(f) show the evolution of a quasiperiodic attractor into a strange attractor as the intensity of choice increases. Remaining game parameters: $b = 2.48, c = 1, \varepsilon = 0.01, n = 0.3$. 

(a) $b = 2.48$. Bifurcation diagram $\beta$
(b) $b = 2.48$. Lyapunov exponent
(c) $\beta = 9.05$. 1-piece attractor
(d) $\beta = 10.05$. 6-piece quasiperiodic
(e) $\beta = 10.2$. 6-piece chaotic
(f) $\beta = 10.8$. 1-piece chaotic
5.6.1 Numerical Bifurcation Curves

The visual analysis of the bifurcation diagram in Fig. 5.25 can be confirmed by a rigorous continuation procedure, i.e. the computation of curves of equilibria along with their detected codimension I bifurcations as one parameter is varied using the Cl_Matcont bifurcation software package\textsuperscript{22} (Dhooge et al. (2003)). One such fixed point - (0.4972065; 0.4178546; 0.06911757; 0.01521537; 0.01521537) for the game parameterization $b = 1.1; c = 1; \varepsilon = 0.01; n = 0.3; \beta = 14$ - is continued in Fig. 5.28 top left panel, as the benefit of cooperation $b$ is increased. Three singularities are detected by the continuation package in the meaningful region $[b > c = 1]$ of $b$ : two Neimark-Sacker (NS) points at $b = 1.270046$ and $b = 2.636476$\textsuperscript{23} and one Period-Doubling (PD) point at $b = 3.158022$. These two points are next 'continued' with respect to another parameter as depicted with codimension II singularities\textsuperscript{24} detected along them. The bifurcation curves describe in a systematic way the qualitative changes in the behavior of the dynamical system in a certain parameters space. For instance, in the two bottom panels in Fig. 5.28 the plotted curves of Neimark-Sacker singularities partition the parameters space in regions with qualitatively similar behavior: when crossing a NS curve from below the system loses stability of the steady state and a stable limit cycles arises around the unstable steady state. Such an analysis reveals that the behavior envisaged by simulations is robust to perturbations of the repeated Prisoner’s Dilemma game (as parametrized by the $b/c$ ratio), the probability of mistake in implementing a particular strategy $\varepsilon$, and the generosity parameter $n$. Note that, compared to the $3 \times 3$ ecologies, there is no fold bifurcation as coordination on equilibria is more difficult to achieve with more strategies. Instead, increasing the strategy space generates oscillatory behavior more easily.

\textsuperscript{22}Cl_Matcont for maps, see package documentation at http://www.matcont.ugent.be/

\textsuperscript{23}normal form coefficients are -2.741700 and -5.440686 meaning that the NS bifurcations are supercritical, i.e. the limit cycles are born stable.

\textsuperscript{24}e.g. a Chenciner CH codim II bifurcation gives rise to much more complicated patterns of behavior (see Kuznetsov (1995) pp. 482).
Figure 5.28: AllD vs. TFT vs. GTFT vs. WSLS vs. AllC. Top left Panel continues a numerically computed fixed point, in the benefit of cooperation space. Detected codimension I singularities are then continued with respect to another game/behavioral parameter. The resulting curves of Period-Doubling (PD) and Neimark-Sacker (NS) points, with respect to the intensity of choice $\beta$, probability of error in strategy execution $\varepsilon$, and degree of generosity $n$, are plotted in the top-right, middle and bottom panels, respectively. The occurrence of codimension II singularities can be observed along all these curves. Unless free to float, game parameters set to: $b = 2$, $c = 1$, $\varepsilon = 0.01$, $n = 0.3$, $\beta = 15$
5.7 Conclusions

In an evolutionary set-up, we append an ecology of iterated Prisoner’s Dilemma (IPD) game strategies, consisting of unconditional cooperators (AllC), unconditional defectors (AllD) and reactive players (TFT) with two repeated strategies that have received less attention in the evolutionary IPD game literature, the error-proof, "generous" tit-for-tat (GTFT) which, with a certain probability, re-establishes cooperation after a (possibly by mistake) defection of the opponent and the penitent, "stimulus-response" (WSLS) strategy that resets cooperation after the opponent punished for defection. Stable oscillations in the frequency of both the forgiving (GTFT) and repentant (Pavlov) strategy along with chaotic behavior emerge under a perturbed version of best-response dynamics, the logit dynamics. We have performed a detailed analysis of the relatively simple $2 \times 2$ and $3 \times 3$ case with the dynamic behaviour now fairly well understood. However, the enlarged $4 \times 4$ and $5 \times 5$ exhibit complicated behavior via period-doubling and breaking of an invariant circle routes to chaos. Finite intensity of choice $\beta$ leads to rich dynamics, path-dependence and co-existence of cyclical and chaotic attractors for a wide selection of $4 \times 4$ ecologies. If we turn to the best-reply limit of the logit dynamics, various 3—and 4—cycles are exhibited in specific $4 \times 4$ and in the $5 \times 5$ ecologies. Last, in terms of individual strategies performance, we discovered first that the presence of unconditional cooperators turns out detrimental to the discriminating types (TFT, GTFT and WSLS) in some $4 \times 4$ interactions leading the population to an AllD monomorphism. Second, there is mixed evidence for a Pavlovian meta-rule in an evolutionary repeated PD with a population of rational players: on the one hand, the stimulus-response strategy does very well and wins the evolutionary competition in a $4 \times 4$ environment with hard defectors (AllD) but no undiscriminating cooperators (AllC), but it almost goes extinct in the full $5 \times 5$ repeated game. Nevertheless, with boundedly rational players, the fractions of Pavlov maintain high values even within the complete ecology of five rules.
5.A Iterated PD Game-stationary distributions and average payoffs

For each pairwise interaction of the iterated PD strategies $S_i \times S_j$, the stationary distribution $\Omega = \{\tau_{CC}, \tau_{CD}, \tau_{DC}, \tau_{DD}\}$ gives the average fraction of time system spends in each of the four states $CC, CD, DC, DD$, respectively. Average expected payoffs are computed as weighted average of the stage game matrix payoffs with weights obtained from the stationary distribution:

\[
\pi_i(S_i, S_j) = \begin{bmatrix}
    b - c \\
    -c \\
    b \\
    0
\end{bmatrix}^T
\begin{bmatrix}
    \tau_{CC} \\
    \tau_{CD} \\
    \tau_{DC} \\
    \tau_{DD}
\end{bmatrix}
\]

\[i, j \in \{\text{AllD, TFT, GTFT, WSL}, \text{AllC}\}\]

Using the transition probability matrix (5.2) we can derive the stationary distribution together with the respective average expected payoffs as summarized in the following table (some expression $m_{ij}$ are too long and given at the end of the table):
<table>
<thead>
<tr>
<th>Interaction</th>
<th>Stationary Distribution</th>
<th>Average Payoff (^{25})</th>
</tr>
</thead>
</table>
| AllD vs. AllD | \(\varepsilon^2\)  
  \(-\varepsilon(\varepsilon - 1)\)  
  \(-\varepsilon(\varepsilon - 1)\)  
  \((\varepsilon - 1)^2\) | \(m_{11} = \varepsilon(b - c)\) |
| AllD vs. TFT | \(-2\varepsilon^2(\varepsilon - 1)\)  
  \(2\varepsilon^3 - 2\varepsilon^2 + \varepsilon\)  
  \(2\varepsilon(\varepsilon - 1)^2\)  
  \(-2\varepsilon^3 + 4\varepsilon^2 - 3\varepsilon + 1\) | \(m_{12} = -\varepsilon(c - 2b + 2b\varepsilon)\) |
| AllD vs. GFT | \(\frac{\varepsilon(1-\varepsilon)(n+\varepsilon)}{n+m+m\varepsilon-n\varepsilon+1}\)  
  \(\frac{\varepsilon}{n-m+m\varepsilon-n\varepsilon+1}\)  
  \(\varepsilon^2\)  
  \(\frac{(\varepsilon-1)^2(n+\varepsilon)}{n-m+m\varepsilon-n\varepsilon+1}\)  
  \(\frac{(1-\varepsilon)(n+m\varepsilon-n\varepsilon+1)}{n-m+m\varepsilon-n\varepsilon+1}\) | \(m_{13}\) |
| AllD vs. WSLS | \(\frac{\frac{1}{2}\varepsilon}{\frac{1}{2}+\varepsilon}\)  
  \(\frac{\frac{1}{2}\varepsilon}{\frac{1}{2}+\frac{1}{2}\varepsilon}\)  
  \(\frac{\frac{1}{2} - \frac{1}{2}\varepsilon}{\frac{1}{2}+\frac{1}{2}\varepsilon}\)  
  \(\frac{\frac{1}{2} - \frac{1}{2}\varepsilon}{\frac{1}{2}+\frac{1}{2}\varepsilon}\) | \(m_{14} = \frac{1}{2}b - c\varepsilon\) |
| AllD vs. AHC | \(-\varepsilon(\varepsilon - 1)\)  
  \(\varepsilon^2\)  
  \((\varepsilon - 1)^2\)  
  \(-\varepsilon(\varepsilon - 1)\) | \(m_{15} = b - b\varepsilon - c\varepsilon\) |

\(^{25}\)Stationary distributions are valid for the general formulation of \(GTFT - (1, m, 1, n)\). However, expressions for the average payoffs become very complicated and are therefore computed under the restriction \(m = n\) in the general form of \(GTFT\), with the impact of the asymmetry \((m \neq n)\) in the two probabilities of restoring cooperation (i.e. after a \(CD\) or \(DD\) history) left for further research.
<table>
<thead>
<tr>
<th>Interaction</th>
<th>Stationary Distribution</th>
<th>Payoffs</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFT vs ABD</td>
<td>(-2\varepsilon^2 (\varepsilon - 1))</td>
<td>(m_{21})</td>
</tr>
<tr>
<td></td>
<td>(2\varepsilon (\varepsilon - 1)^2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2\varepsilon^3 - 2\varepsilon^2 + \varepsilon)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2\varepsilon^3 + 4\varepsilon^2 - 3\varepsilon + 1)</td>
<td></td>
</tr>
<tr>
<td>TFT vs TFT</td>
<td>(1/4)</td>
<td>(m_{22} = \frac{1}{2}b - \frac{1}{2}c)</td>
</tr>
<tr>
<td></td>
<td>(1/4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1/4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1/4)</td>
<td></td>
</tr>
<tr>
<td>TFT vs GTFT</td>
<td>(\frac{(1-\varepsilon)(n+\varepsilon)(m+\varepsilon-2m\varepsilon)}{2\varepsilon^2} (\varepsilon - 1)^2 (m+\varepsilon - 2)(\varepsilon + n-2n\varepsilon))</td>
<td>(m_{23})</td>
</tr>
<tr>
<td></td>
<td>(\frac{m\varepsilon+5nc+9c^2-12c^3+4c^4-6mc^2+4mc^3-10nc^2+4nc^4+mn-4mnc+4mnc^2}{2\varepsilon^2\varepsilon-1\varepsilon} (m+\varepsilon-2)(\varepsilon + n-2n\varepsilon))</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\frac{m\varepsilon+5nc+9c^2-12c^3+4c^4-6mc^2+4mc^3-10nc^2+4nc^4+mn-4mnc+4mnc^2}{2\varepsilon^2\varepsilon-1\varepsilon} (m+\varepsilon-2)(\varepsilon + n-2n\varepsilon))</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\frac{m\varepsilon+5nc+9c^2-12c^3+4c^4-6mc^2+4mc^3-10nc^2+4nc^4+mn-4mnc+4mnc^2}{2\varepsilon^2\varepsilon-1\varepsilon} (m+\varepsilon-2)(\varepsilon + n-2n\varepsilon))</td>
<td></td>
</tr>
<tr>
<td>TFT vs WSLS</td>
<td>(1/4)</td>
<td>(m_{24} = \frac{1}{2}b - \frac{1}{2}c)</td>
</tr>
<tr>
<td></td>
<td>(1/4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1/4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1/4)</td>
<td></td>
</tr>
<tr>
<td>TFT vs AIC</td>
<td>(- (\varepsilon - 1) (2\varepsilon^2 - 2\varepsilon + 1))</td>
<td>(m_{25})</td>
</tr>
<tr>
<td></td>
<td>(-2\varepsilon^2 \varepsilon=1\varepsilon -2\varepsilon^2 (2\varepsilon^2 - 2\varepsilon + 1))</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2\varepsilon (\varepsilon - 1)^2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2\varepsilon^2 (\varepsilon - 1))</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>Stationary Distribution</td>
<td>Payoffs</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>GTFT vs. AllD</td>
<td>(-\varepsilon (\varepsilon - 1) \frac{n+\varepsilon}{n-m+me-\varepsilon n+1}) \quad (\varepsilon - 1)^2 \frac{n+\varepsilon}{n-m+me-\varepsilon n+1}) \quad \varepsilon \frac{mn-\varepsilon}{n-m+me-\varepsilon n+1}) \quad \varepsilon^2 \frac{m+\varepsilon}{n-m+me-\varepsilon n+1})</td>
<td>(m_{31})</td>
</tr>
<tr>
<td>GTFT vs. TFT</td>
<td>(\frac{(1-\varepsilon)(n+\varepsilon)(m+\varepsilon-2n\varepsilon)}{2\varepsilon(\varepsilon-1)^2(n+\varepsilon)}) \quad \frac{me+5ne+9\varepsilon^2-12\varepsilon^3+4\varepsilon^4-6n\varepsilon^2+4m\varepsilon^3-10\varepsilon n^2+4n\varepsilon^3+mn-4mn^2+4m^2n^2}{2\varepsilon(\varepsilon-1)^2(n+\varepsilon)})</td>
<td>(m_{32})</td>
</tr>
<tr>
<td>GTFT vs. GTFT</td>
<td>(-\varepsilon (\varepsilon - 1) \frac{2n-2m\varepsilon+ne-\varepsilon mn}{n+\varepsilon-2}) \quad \frac{n\varepsilon}{m+\varepsilon-2}) \quad \frac{n\varepsilon}{m+\varepsilon-2}) \quad \frac{\varepsilon}{m+\varepsilon-2}) \quad \frac{\varepsilon^2}{m+\varepsilon-2}) \quad \frac{n^2\varepsilon-2mn+2n\varepsilon}{m+\varepsilon-2}) \quad \frac{mn^2+2m^2n+2mn^2+2mn^2+2n^2\varepsilon-2mn+2n\varepsilon}{m+\varepsilon-2}) \quad \frac{m^2n+2mn^2+2mn^2+2mn^2+2n^2\varepsilon-2mn+2n\varepsilon}{m+\varepsilon-2})</td>
<td>(m_{33})</td>
</tr>
<tr>
<td>GTFT vs. WLS</td>
<td>(\frac{(n+\varepsilon-2n\varepsilon+\varepsilon^2+2n^2\varepsilon-mn+2mn^2)(\varepsilon-1)}{4me-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8n\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn^2+4m^2n^2}{(n-2\varepsilon+\varepsilon^2+1)^2(\varepsilon-1)}) \quad \frac{4me-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8n\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn^2+4m^2n^2}{(n-2\varepsilon+\varepsilon^2+1)^2(\varepsilon-1)}) \quad \frac{-2(2\varepsilon-2m\varepsilon-3\varepsilon^2+4n\varepsilon^3+2m^2n+mn-2mn^2)(\varepsilon-1)}{4me-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8n\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn^2+4m^2n^2}{(n-2\varepsilon+\varepsilon^2+1)^2(\varepsilon-1)}) \quad \frac{n\varepsilon}{n-\varepsilon+\varepsilon^2+1}) \quad \frac{4me-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8n\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn^2+4m^2n^2}{(n-2\varepsilon+\varepsilon^2+1)^2(\varepsilon-1)})</td>
<td>(m_{34})</td>
</tr>
<tr>
<td>GTFT vs. AIC</td>
<td>(-\varepsilon \frac{n\varepsilon-2\varepsilon+\varepsilon^2+1}{n\varepsilon-m\varepsilon+1}) \quad \frac{\varepsilon (\varepsilon - 1) \frac{m+\varepsilon-2}{n\varepsilon-m\varepsilon+1}}{\varepsilon^2 \frac{m+\varepsilon}{n\varepsilon-m\varepsilon+1}})</td>
<td>(m_{35})</td>
</tr>
<tr>
<td>Interaction</td>
<td>Stationary Distribution</td>
<td>Payoffs</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>W SLS vs. AlID</td>
<td>$\begin{align*} \frac{1}{2} &amp; \varepsilon \ \frac{1}{2} - \frac{1}{2} &amp; \varepsilon \ \frac{1}{2} &amp; \varepsilon \ \frac{1}{2} - \frac{1}{2} &amp; \varepsilon \end{align*}$</td>
<td>$m_{41} = b\varepsilon - \frac{1}{2}c$</td>
</tr>
<tr>
<td>W SLS vs. TFT</td>
<td>$\begin{align*} 1 &amp; /4 \ 1 &amp; /4 \ 1 &amp; /4 \ 1 &amp; /4 \end{align*}$</td>
<td>$m_{42} = \frac{1}{2}(b - c)$</td>
</tr>
<tr>
<td>W SLS vs. GTFT</td>
<td>$\begin{align*} (n + \varepsilon - 2n\varepsilon - \varepsilon^2 + 2n\varepsilon^2 - mn + 2mn\varepsilon)(\varepsilon - 1) \ 4m\varepsilon - 7n - 4n\varepsilon + 15\varepsilon^2 - 12\varepsilon^3 + 4\varepsilon^4 - 8n\varepsilon^2 + 4m\varepsilon^3 - 8n\varepsilon^2 + 4n\varepsilon^3 + mn - 4m\varepsilon + 4mn\varepsilon + 4mn\varepsilon^2 \ -(2n - 2m - n\varepsilon - 3\varepsilon^2 + 3\varepsilon^3 + 4n\varepsilon^2 - 2m\varepsilon^3 + 2n\varepsilon^2 + mn\varepsilon - 2mn\varepsilon^2) \ 4m\varepsilon - 7n - 4n\varepsilon + 15\varepsilon^2 - 12\varepsilon^3 + 4\varepsilon^4 - 8n\varepsilon^2 + 4m\varepsilon^3 - 8n\varepsilon^2 + 4n\varepsilon^3 + mn - 4m\varepsilon + 4mn\varepsilon + 4mn\varepsilon^2 \ (n\varepsilon - 2n\varepsilon + \varepsilon^2 + 1)2\varepsilon(\varepsilon - 1) \ 4m\varepsilon - 7n - 4n\varepsilon + 15\varepsilon^2 - 12\varepsilon^3 + 4\varepsilon^4 - 8n\varepsilon^2 + 4m\varepsilon^3 - 8n\varepsilon^2 + 4n\varepsilon^3 + mn - 4m\varepsilon + 4mn\varepsilon + 4mn\varepsilon^2 \end{align*}$</td>
<td>$m_{43}$</td>
</tr>
<tr>
<td>W SLS vs. W SLS</td>
<td>$\begin{align*} -4\varepsilon^3 + 7\varepsilon^2 - 4\varepsilon + 1 \ -\varepsilon (\varepsilon - 1) \ -\varepsilon (\varepsilon - 1) \ 4\varepsilon^3 - 5\varepsilon^2 + 2\varepsilon \end{align*}$</td>
<td>$m_{44}$</td>
</tr>
<tr>
<td>W SLS vs. AlC</td>
<td>$\begin{align*} \frac{1}{2} - \frac{1}{2} \varepsilon \ \frac{1}{2} \varepsilon \ \frac{1}{2} - \frac{1}{2} \varepsilon \ \frac{1}{2} \varepsilon \end{align*}$</td>
<td>$m_{45}$</td>
</tr>
<tr>
<td>Interaction</td>
<td>Stationary Distribution</td>
<td>Payoffs</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>AllC vs. AID</td>
<td>$-\varepsilon (\varepsilon - 1)$</td>
<td>$m_{51} = b\varepsilon - c + c\varepsilon$</td>
</tr>
<tr>
<td></td>
<td>$(\varepsilon - 1)^2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\varepsilon^2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-\varepsilon (\varepsilon - 1)$</td>
<td></td>
</tr>
<tr>
<td>AllC vs. TFT</td>
<td>$-2\varepsilon^3 + 4\varepsilon^2 - 3\varepsilon + 1$</td>
<td>$m_{52} = b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2$</td>
</tr>
<tr>
<td></td>
<td>$2\varepsilon (\varepsilon - 1)^2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$2\varepsilon^3 - 2\varepsilon^2 + \varepsilon$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-2\varepsilon^2 (\varepsilon - 1)$</td>
<td></td>
</tr>
<tr>
<td>AllC vs. GTFT</td>
<td>$-\frac{\varepsilon - 1}{n\varepsilon - m\varepsilon + 1} (n\varepsilon - 2\varepsilon + \varepsilon^2 + 1)$</td>
<td>$m_{53} = b - c - 2b\varepsilon + c\varepsilon + b\varepsilon^2 + b\varepsilon$</td>
</tr>
<tr>
<td></td>
<td>$\varepsilon (\varepsilon - 1) \frac{m + \varepsilon - 2}{n\varepsilon - m\varepsilon + 1}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\frac{\varepsilon}{n\varepsilon - m\varepsilon + 1} (n\varepsilon - 2\varepsilon + \varepsilon^2 + 1)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-\varepsilon^2 \frac{m + \varepsilon - 2}{n\varepsilon - m\varepsilon + 1}$</td>
<td></td>
</tr>
<tr>
<td>AllC vs. WSLS</td>
<td>$-2\varepsilon^3 + 4\varepsilon^2 - 3\varepsilon + 1$</td>
<td>$m_{54} = b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2$</td>
</tr>
<tr>
<td></td>
<td>$2\varepsilon (\varepsilon - 1)^2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$2\varepsilon^3 - 2\varepsilon^2 + \varepsilon$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-2\varepsilon^2 (\varepsilon - 1)$</td>
<td></td>
</tr>
<tr>
<td>AllC vs. AllC</td>
<td>$(\varepsilon - 1)^2$</td>
<td>$m_{55} = (1 - \varepsilon) (b - c)$</td>
</tr>
<tr>
<td></td>
<td>$-\varepsilon (\varepsilon - 1)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-\varepsilon (\varepsilon - 1)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\varepsilon^2$</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.4: Stationary distributions and average expected payoffs for an iterated PD game with an ecology of repeated rules consisting of AllD, TFT, GTFT, WSLS and AllC.
where,

\[ m_{13} = b \varepsilon - c \varepsilon - b \varepsilon^2 + bn - bn \varepsilon \]

\[ m_{21} = \varepsilon (b - 2c + 2c \varepsilon) \]

\[ m_{23} = \frac{-\varepsilon - b \varepsilon^2 + b n + c n + b n \varepsilon - 2 c n \varepsilon}{n + 3 \varepsilon - 2 n \varepsilon - 2 \varepsilon^2} \]

\[ m_{25} = b - c - b \varepsilon + 2 c \varepsilon - 2 c \varepsilon^2 \]

\[ m_{31} = b \varepsilon - c \varepsilon + c \varepsilon^2 - cn + cn \varepsilon \]

\[ m_{32} = \frac{b \varepsilon - c \varepsilon + c \varepsilon^2 + b n - c n - 2 b n \varepsilon + c n}{n + 3 \varepsilon - 2 n \varepsilon - 2 \varepsilon^2} \]

\[ m_{34} = -m_{341} / m_{342}, \text{ with } \]

\[ m_{341} = 3 b \varepsilon - 3 c \varepsilon - b n^2 + c n^2 - 5 b \varepsilon^2 + 2 b \varepsilon^3 + 8 c \varepsilon^2 - 7 c \varepsilon^3 + 2 c \varepsilon^4 + b n - c n - 4 b n^2 \varepsilon^2 + 2 c n^2 \varepsilon^2 - 6 b n \varepsilon + 3 c n \varepsilon + 10 b n \varepsilon^2 + 4 b n^2 \varepsilon - 4 b n^2 \varepsilon^3 - 6 c n \varepsilon^2 - 3 c n^2 \varepsilon + 4 c n \varepsilon^3 \]

\[ m_{342} = 4 n^2 \varepsilon^2 - 4 n^2 \varepsilon + n^2 + 8 n \varepsilon^3 - 16 n \varepsilon^2 + 8 n \varepsilon - n + 4 \varepsilon^4 - 12 \varepsilon^3 + 15 \varepsilon^2 - 7 \varepsilon \]

\[ m_{35} = b - c - b \varepsilon + 2 c \varepsilon - c \varepsilon^2 - c n \varepsilon \]

\[ m_{43} = m_{431} / m_{432}, \text{ with } \]

\[ m_{431} = 3 c \varepsilon - 3 b \varepsilon + b n^2 - c n^2 + 8 b \varepsilon^2 - 7 b \varepsilon^3 + 2 b \varepsilon^4 - 5 c \varepsilon^2 + 2 c \varepsilon^3 - b n + c n + 2 b n^2 \varepsilon^2 - 4 c n^2 \varepsilon^2 + 3 b n \varepsilon - 6 c n \varepsilon - 6 b n \varepsilon^2 - 3 b n^2 \varepsilon + 4 b n \varepsilon^3 + 10 c n \varepsilon^2 + 4 c n^2 \varepsilon - 4 c n^3 \varepsilon \]

\[ m_{44} = (b - c) (1 - 4 \varepsilon^3 + 6 \varepsilon^2 - 3 \varepsilon) \]

\[ m_{45} = b - \frac{1}{2} c - b \varepsilon \]
Chapter 6

Summary

The aim of this thesis is to study the dynamics of evolutionary games. Initially, evolutionary game theory has been viewed as an evolutionary "repair" of rational actor game theory in the hope that a population of boundedly rational players may attain convergence to classic "rational" solutions such as the Nash Equilibrium through some learning or evolutionary process occurring at the population level. However, there is still no class of evolutionary dynamics, neither simple, imitative learning (Replicator Dynamics) nor involved, belief-based learning (Fictitious Play or Best Reply Dynamics), guaranteeing convergence to Nash equilibrium, beyond certain restricted classes of games, usually with a small number of strategies.

In this thesis the model of boundedly rational players is a perturbed version of the best-reply choice, the so-called Logit rule. The strategic context to which this evolutionary process is applied varies from simple models of cyclical competition (Rock-Paper-Scissors), through industrial organization (Cournot game) and to collective-action choice (repeated Prisoner's Dilemma) games. The logit choice rule has the appealing property of capturing a continuum of degrees of (myopic) rational behavior through differential responsiveness to payoff differences between alternative choices. The connecting line through the chapters lies in the question whether a steady state is obtained as long run outcome of the Logit evolutionary dynamics or
whether multiple steady states (path-dependence) or more complicated, cyclical or perhaps chaotic attractors, may emerge as well.

Chapter 2 contrasts the long-run behavior of Replicator and Logit Dynamics in standard Rock-Paper-Scissors and Coordination games. It is well known that Replicator Dynamics does not give rise to a stable limit cycle but to heteroclinic cycles in $3 \times 3$ circulant RSP games (Zeeman (1980)). We obtain an alternative proof of this classical result based on the computation of the first Lyapunov coefficient in the normal form of the vector field induced by the Replicator Dynamics, and show that all Hopf bifurcations are degenerate. Via the same technique, generic Hopf bifurcations are shown to occur under the Logit Dynamics and, moreover, all these bifurcations are supercritical, i.e. the system exhibits stable limit cycles. This result is extended, through continuation analysis with the Matcont software package to the class of $3 \times 3$ non-circulant RSP games. Besides generic Hopf bifurcation, fold singularities (i.e. creation of multiple interior steady states) also do not occur in $3 \times 3$ games under Replicator Dynamics (Zeeman (1980)). With the help of a simple $3 \times 3$ Coordination game we detect numerically a scenario with three consecutive fold bifurcations under the Logit Dynamics giving rise to multiple interior steady states. Both the Hopf and fold codimension I singularities are continued in the game payoffs parameters and a behavioral parameter such as the intensity of choice, in order to obtain bifurcation curves in the parameter space along with certain codimension II singularities along these curves. A frequency-weighted version of the Logit dynamics (iLogit) is put at work on the circulant $3 \times 3$ RSP game and shown to generate stable limit cycles with a large amplitude reminiscent of the heteroclinic cycle connecting the three monomorphic steady states in the Replicator Dynamics. Unlike Logit Dynamics, iLogit runs into chaotic behavior in an example of 4x4 symmetric, biologically-inspired game.

While Chapter 2 focuses on the actions chosen in a particular strategic situation, Chapter 3 investigates a deeper level of evolution of learning rules, in the context of a
linear inverse demand-quadratic costs Cournot duopoly. Besides choosing an action in the stage quantity-setting game as a response to expectations about the choice of the opponent (i.e. her quantity produced), players also have to make a decision about the learning heuristic employed to form their expectations. This higher-order competition between the expectation-formation rules has an impact on the stability property of the underlying stage game Cournot-Nash equilibrium. Various pairs of ecologies are drawn from a toolbox consisting of naive, adaptive, (weighted) fictitious play or rational (Nash) strategy. For instance, the evolutionary competition between a free adaptive and a costly weighted fictitious play rule destabilizes the Cournot-Nash equilibrium through a primary period-doubling bifurcation and a secondary Neimark-Sacker bifurcation and eventually enters a chaotic regime. This analysis is extended in Chapter 4 to linear inverse demand linear costs Cournot games with an arbitrary number of players in the spirit of the Theocharis (1960) "two is stable, three is unstable" result. In a heterogenous learning rules environment consisting of adaptive and rational players we show that the Cournot-Nash equilibrium destabilizes through a period-doubling route to chaos as the number of players increases. Theocharis (1960) unstable triopoly is derived as the limit of a system with homogenous (no switching) naive expectation. Interestingly, it could be stabilized by adjusting the model parameters, for instance, by making the expectations more adaptive (placing higher weights on remote past observations).

Bounded rationality models with heterogenous players endowed with smart and simple heuristics (Gigerenzer and Todd (1999)) emerges as a promising way out of the conundrum of observed cooperation in collective action dilemmas. An ecology of simple rules in an iterated Prisoner’s Dilemma framework is constructed in Chapter 5 with the selection of rules driven by a logistic-type of evolutionary dynamics. It consists of reciprocators, unconditional defectors and cooperators, stimulus-response (Pavlov) and generous reciprocators all subject to small implementation errors in strategy execution and small mistakes in selecting the best-performing heuristic
available in the population. The evolutionary selection dynamics in the resulting subecologies of rules are investigated by a mixture of analytical and numerical methods, with rich dynamical behaviour unfolding. The surrounding ecology appears critical for the success or failure of certain heuristics and there is no "undisputed" winner. Pavlov (stimulus-response) players do well in an environment without unconditional cooperators, but poorly when such indiscriminately "nice" players are around. Rock-Paper-Scissors cycles arise in the $3 \times 3$ ecologies while some $4 \times 4$ ecologies easily display path-dependence and co-existence of periodic and chaotic attractors.

A general conclusion of this thesis is that the evolutionary selection among boundedly rational strategies does not necessarily ensure convergence to Nash equilibrium and a richer dynamical behavior may be the rule rather than the exception.
Bibliography


Samenvatting

Het doel van dit proefschrift is om de dynamica van evolutionaire spellen te be-
studeren. Aanvankelijk werd evolutionaire speltheorie gezien als een manier om
toantele speltheorie te “repareren”. De hoop was dat het geëggregeerde gedrag
van een populatie van begrensd toantele spelers convergeert naar klassieke toantele
oplossingsconcepten zoals het Nash-evenwicht door middel van leer- of evolution-
aire processen. Er is echter nog steeds geen klasse van evolutionaire dynamica,
nog eenvoudige imitatie (replicatordynamica), noch gecompliceerd leergedrag op
basis van verwachtingen over het gedrag van andere spelers (fictief spelen of beste-
toantwoorddynamica), die garanderen dat er convergentie is naar het Nash-evenwicht
tenzij er gekeken wordt naar een beperkte klasse van spelen met meestal een klein
aantal strategieën.

In dit proefschrift worden beperkt toantele spelers gemodelleerd met een gep-
erturbeerde versie van de beste-toantwoorddynamica, de zogeheten logitregel. De
strategische context, waarin dit evolutionaire proces wordt toegepast, varieert van
eenvoudige cyclische competitiemodellen (Steen-Papier-Schaar, SPS in het vervolg),
tot bedrijfstakorganisatie (Cournot-oligopolie) en publieke goederenspellen zoals het
herhaalde gevangenendilemma. De logitregel heeft de aantrekkelijke eigenschap dat
het verscheidene gradaties van (bijziende) toanteel gedrag kan omvatten door de re-
gel meer of minder gevoelig te maken voor verschillen in de aantrekkelijkheid van
keuzes. De rode draad door de hoofdstukken is de vraag of een vaste toestand op
de lange termijn de uitkomst is van de evolutionaire dynamica, gekenmerkt door het
gebruik van de logitregel, of dat er gecompliceerder gedrag kan ontstaan. Hierbij kan men denken aan meervoudige evenwichten (padafhankelijkheid), cyclische of zelfs chaotische aantrekkers.

Hoofdstuk 2 vergelijkt het gedrag op de lange termijn van de replicatordynamica met het gedrag van de logitdynamica in standaard SPS- en coördinatiespellen. Het is bekend dat replicatordynamica niet leidt tot stabiele limietcykels maar tot heterocliene cykels in $3 \times 3$ circulaire SPS-spellen (Zeeman (1980)). Wij verkrijgen een alternatief bewijs van dit klassieke resultaat gebaseerd op de berekening van de eerste Lyapunov-coëfficiënt in de normaalvorm van het vectorveld geïnduceerd door de replicatordynamica en laten zien dat alle Hopf-bifurcaties gedegeneerd zijn. Met dezelfde techniek laten we zien dat generieke Hopf-bifurcaties wel voorkomen in de logitdynamica en zelfs dat deze bifurcaties superkritiek zijn, d.w.z. het systeem vertoont stabiele limietcykels. Dit resultaat wordt uitgebreid door middel van continuatietechnieken en met behulp van het softwarepakket Matcont naar de klasse van $3 \times 3$ niet-circulaire SPS-spellen. Net als generieke Hopf-bifurcaties komen ook zadelknopen niet voor in $3 \times 3$ spellen onder replicatordynamica (Zeeman (1980)). Met behulp van een eenvoudig $3 \times 3$ coördinatiespel vinden wij numeriek een scenario waarin drie opeenvolgende zadelknoopbifurcaties onder logitdynamica leiden tot meervoudige inwendige vaste toestanden. Zowel de Hopf-bifurcatie als de codimensie I zadelknoop worden gecontinueerd in de payoffparameters van het spel en gedragsparameters zoals de keuzeintensiteit. Hieruit volgen bifurcatiekrommen in de parameterruimte waarlangs codimensie II singulariteiten zich bevinden. Vervolgens bestuderen wij een frequentie-gewogen versie van de logitdynamica (wlogit), die wij toepassen op een circulair $3 \times 3$ SPS-spel. Wij laten zien dat stabiele limietcykels met een grote amplitude voorkomen. Deze cykel vertoont overeenkomsten met de heterocliene cykel die de drie eenvormige vaste toestanden onder de replicatordynamica verbind. Anders dan logitdynamica, kan wlogit chaotisch gedrag vertonen in een symmetrisch $4 \times 4$-spel. Het voorbeeld, dat wij construeren, is geïnspireerd door
de biologie.


Modellen met begrensd rationele en heterogene spelers uitgerust met slimme én eenvoudige heuristieken (Gigerenzer and Todd (1999)) lijken een veelbelovende manier om het raadsel van samenwerking in het gevangendilemma op te lossen. Een
ecologie van eenvoudige regels in een geëntereerd gevangendilemmaanwerk wordt geconstrueerd in Hoofdstuk 5 en de selectie van regels wordt aangedreven door logitachtige evolutionaire dynamica. Deze ecologie bestaat uit reciprocators, onconditionele afwijkers, onconditionele samenwerkers, Pavlov (stimulus-respons) en vergevingsgezinde reciprocators, die allemaal met een kleine kans een fout kunnen maken bij het uitvoeren van hun actie en een kleine fout in het kiezen van de beste heuristiek die beschikbaar is. Met een mengeling van analytische en numerieke methoden wordt de rijke evolutionaire dynamica van de resulterende deelecologie bestudeerd. De ecologie is cruciaal voor het succes en het mislukken van bepaalde heuristieken en er is geen duidelijk beste heuristiek. Pavlov (stimulus-respons) spelers doen het goed in omgeving waarin geen onconditionele samenwerkers zijn, maar varen slecht bij de aanwezigheid van spelers die altijd “aardig” zijn. SPS-achtige cykels ontstaan in $3 \times 3$ ecologieën, terwijl sommige $4 \times 4$ ecologieën gevoelig zijn voor padafhankelijkheid en coëxistentie van periodieke en chaotische aantrekkers.

Een algemene conclusie van dit proefschrift is dat evolutionaire selectie tussen begrensd rationele strategieën niet noodzakelijk convergentie naar het Nash-evenwicht impliceert en dat rijk dynamisch gedrag eerder regel dan uitzondering is.
The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

419. L. RATNOVSKI, A Random Walk Down the Lombard Street: Essays on Banking.
420. R.P. NICOLAI, Maintenance models for systems subject to measurable deterioration.
421. R.K. ANDADARI, Local clusters in global value chains, a case study of wood furniture clusters in Central Java (Indonesia).
422. V. KARTSEVA, Designing Controls for Network Organizations: A Value-Based Approach.
423. J. ARTS, Essays on New Product Adoption and Diffusion.
425. M. VAN DER VOORT, Modelling Credit Derivatives.
427. E. BEKKERS, Essays on Firm Heterogeneity and Quality in International Trade.
430. W. VERMEULEN, Essays on Housing Supply, Land Use Regulation and Regional Labour Markets.
437. R. LORD, Efficient pricing algorithms for exotic derivatives.
439. Y.-Y. TSENG, Valuation of travel time reliability in passenger transport.
440. M.C. NON, Essays on Consumer Search and Interlocking Directorates.
441. M. DE HAAN, Family Background and Children’s Schooling Outcomes.
442. T. ZAVADIL, Dynamic Econometric Analysis of Insurance Markets with Imperfect Information.
443. I.A. MAZZA, Essays on endogenous economic policy.
444. R. HAIJEMA, Solving large structured Markov Decision Problems for perishable-inventory management and traffic control.
446. R. SEGERS, Advances in Monitoring the Economy.
448. L. PAN, Poverty, Risk and Insurance: Evidence from Ethiopia and Yemen.

200
454. S. VUJIĆ, Econometric Studies to the Economic and Social Factors of Crime.
455. F. HEUKELOM, Kahneman and Tversky and the Making of Behavioral Economics.
460. O.E. JONKEREN, Adaptation to Climate Change in Inland Waterway Transport.
462. J. NIEMCZYK, Consequences and Detection of Invalid Exogeneity Conditions.
463. I. BOS, Incomplete Cartels and Antitrust Policy: Incidence and Detection
464. M. KRAWCZYK, Affect and risk in social interactions and individual decision-making.
465. T. C. LIN, Three Essays on Empirical Asset Pricing
466. J.A. BOLH AAR, Health Insurance: Selection, Incentives and Search
467. T. FARENHORST-YUAN, Efficient Simulation Algorithms for Optimization of Discrete Event Based on Measure Valued Differentiation
468. M. I. OCHEA, Essays on Nonlinear Evolutionary Game Dynamics