



UvA-DARE (Digital Academic Repository)

Essays on nonlinear evolutionary game dynamics

Ochea, M.I.

Publication date
2010

[Link to publication](#)

Citation for published version (APA):

Ochea, M. I. (2010). *Essays on nonlinear evolutionary game dynamics*. [Thesis, fully internal, Universiteit van Amsterdam]. Thela Thesis.

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: <https://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.

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 educative 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 choice² 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 :

²Intensity of choice defines players' sensitivity to payoffs differences between strategies. It may vary from 0 (complete insensitivity to payoff differentials, or random choice) to $+\infty$ (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 expectation³ 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-time⁴: 2-person zero-sum (Robinson (1951)), dominance solvable (Milgrom and Roberts (1991), even for more general classes of adaptive and sophisticated learning), 2-person 2×2 nondegenerate⁵ (Miyasawa (1961)), 2-person $2 \times n$ (Berger (2005)), supermodular with diminishing returns⁶ (Krishna (1992)), 3×3 supermodular (Hahn (1999)), $3 \times n$ and 4×4 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-

³A 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)).

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

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

⁶Games with strategic complementarities. Proof is based on a particular tie-breaking rule. Berger (2008) proves the result for ordinal strategic complementarities 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 stationarity⁷, 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 3×3 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 Dynamics⁸ in 3×3 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 3×3 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 inductive⁹ 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-

⁷All Nash equilibria of the underlying games are rest point of the dynamics.

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

⁹Inductive 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 3×3 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 orbits¹⁰ and repelling anti-Shapley (anticlockwise orbit) polygon. The clockwise and anticlockwise orbits exchange stability as the game payoff parameter is changed. Jordan (1993) discusses a 3-player 2×2 matching penny game and show that the FPP does not hold. Foster and Young (1998) illustrate, by example of a 6×6 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 3×3 games cannot yield isolated period orbits under Replicator Dynamics. The situation changes drastically when one analyses *asymmetric* 3×3 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 3×3 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 4×4 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 behavior¹¹ are exhibited in a 4×4 symmetric game constructed by Berger and Hofbauer

¹⁰Of "jitter type", i.e. neither attracting, repelling or of saddle type

¹¹Survival of strictly dominated strategies.

(2006) under an evolutionary dynamic that inherits features from both replicator and best-reply dynamics: the Brown-von Neumann-Nash dynamics. Chawanya (1997), Schuster et al. (1991), Skyrms (1992) provide analytical and numerical evidence for chaos under Replicator Dynamics for various 4×4 game payoff matrices.

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)¹² study 2-population 2×2 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 2×2 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

¹²To 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 2×2 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 heterogeneous learning, i.e. switching between alternative heuristics.

Chapter 2 investigates the behaviour of simple RSP and 3×3 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. Heterogeneous 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 3×3 stable¹³ 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 3×3 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 3×3 symmetric games: Rock-Paper-Scissors and Coordination game. For the circulant¹⁴ RSP

¹³Zeeman (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".

¹⁴The 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 3×3 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 Matcont¹⁵ (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 3×3 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 welfare¹⁶ 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-

¹⁵A continuation software for ode (see package documentation at <http://www.matcont.ugent.be/>)

¹⁶The 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 4×4 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 fast¹⁷ (a situation dubbed Type I duopoly by Cox and Walker (1998)). However, when one or both firms' marginal cost decrease rapidly enough¹⁸ (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 2×2 rules sub-ecologies Our focus is on the interplay between different learning heuristics and on the impact

¹⁷The two firms' reaction function intersect in a normal way, i.e. they have only one (interior) intersection.

¹⁸Such 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 2×2 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*), indiscriminating defectors (*AllD*), and cooperators (*AllC*) as in Sigmund and Brandt (2006), extended with the Pavlovian "win-stay-lose-shift" (*WSLS*) heuristic¹⁹ and the gen-

¹⁹Pavlov, 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.

erous variation of homo reciprocator "generous tit-for-tat" (*GTFT*)²⁰. The resulting subecologies of 2×2 , 3×3 , 4×4 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 3×3 ecologies comprising Pavlovian and "generous" players, while some 4×4 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* 4×4 environment but poorly when unconditional cooperators are around. However, in the full 5×5 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.

²⁰*GTFT* interprets opponent defection as mistakes and, with certain positive probability, resets cooperation.