BR. Chap 3:

- Team games with single BR (57)
- Coordination games

Unique singleton absorbing = converges to dominant

- Left of risk dominant - Savage basin (3)

Majority equilibrium

BR = Majority

Conditions for no cycle (p.76)

Noise: if plays majority remains

- if not mixed on all

1. boundaries \( \rightarrow \) (ties p.79) \( \rightarrow \)

\((ab) \rightarrow \text{absorbing singleton} \)

Aspiration = majority \( \rightarrow \) coordination singleton

(52) if \( b > 0 \) \( \rightarrow \)

Simulation (39)

Model of prices, heterogenous & absorbing

Theorem 43 (p.119): if only 2nd player act action

Head risk dominant

Extended to risk dominant there

Example: Risk dominant voter

\( \text{not voted for in general} \)

Mistakes depend on payoff (Ref p.122)

\( \text{confident 339} \)
Interaction Patterns, Learning Processes and Equilibria in Population Games

by

Antonella Ianni

Submitted to the on 23 March 1996, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

This thesis concerns the foundations of equilibrium notions in game theory. Game theory and its equilibrium concepts are used today as a tool for the study of many interactive environments, in particular in economics. Recent research focuses on learning processes by means of which a population of players achieves an equilibrium. This thesis contributes to this literature.

The dissertation provides a general framework for the study of interacting settings that involve a finite population of players, and analyzes learning processes that sometimes lead to equilibrium. The idea is that agents learn how to face an interactive contest by reacting adaptively to their environment. In particular, we focus on a population of players, repeatedly and randomly matched to play a symmetric normal form game. We take the view that the more standard assumption of uniform random matching is not always appropriate in an economic setting. We therefore formulate a more general model, where we analyze several simple learning processes. We characterize possible equilibria at the population level, focusing on the relation between the latter and the equilibria of the underlying game. We then examine convergence properties of the processes and interpret attractors of the dynamics in terms of equilibria.

In particular, the thesis analyzes the following issues:

- Models that depart from uniform population matching due to a highly decentralized interaction structure (i.e. locally interactive systems);
- Learning processes where players adopt the following boundedly rational behavioural rules: myopic best-reply, majority rules, (constant) aspiration level learning rules, payoff-independent trembles and payoff-dependent trembles.
- Relation between the equilibria of the population game and equilibria of the
underlying game, and the effects of the interaction structure on the characteri-
zation of dynamic stability of equilibria.
Contents

1 Games and Equilibria ............................................................... 19
  1.1 Introduction ............................................................................. 19
  1.2 The general model ................................................................. 21
    1.2.1 The players: ....................................................................... 21
    1.2.2 The game: ......................................................................... 23
    1.2.3 The population game: ....................................................... 24
  1.3 Equilibria of the game and equilibria of the population game 26
  1.4 Coordination Games ............................................................... 31

2 The Evolution of Play: an Overview ........................................... 35
  2.1 Overview .................................................................................. 35
  2.2 The Evolution of play ............................................................ 38
    2.2.1 Behavioral Rules ............................................................... 39
    2.2.2 Asymptotic behavior ......................................................... 43

3 Myopic Best Reply Dynamics .................................................... 45
  3.1 Introduction ............................................................................. 45
  3.2 The Model ............................................................................... 47
  3.3 The Recurrent Communication Classes. ................................. 50
  3.4 Asymptotic Behaviour and Likelihood of Equilibria. ............... 60

5
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Majority Rules and Noise at the Margin</td>
<td>67</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>67</td>
</tr>
<tr>
<td>4.2</td>
<td>The Model</td>
<td>69</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Local Interaction</td>
<td>69</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Majority Rules and Noise at the Margin</td>
<td>71</td>
</tr>
<tr>
<td>4.3</td>
<td>Optimality of Majority Rules</td>
<td>72</td>
</tr>
<tr>
<td>4.4</td>
<td>The Recurrent Communication Classes</td>
<td>74</td>
</tr>
<tr>
<td>4.5</td>
<td>A Special Structure</td>
<td>77</td>
</tr>
<tr>
<td>5</td>
<td>Aspiration level dynamics: some applications</td>
<td>85</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>85</td>
</tr>
<tr>
<td>5.2</td>
<td>Aspiration Levels</td>
<td>87</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Majority Rules and Payoff memory</td>
<td>89</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Aspiration Levels</td>
<td>93</td>
</tr>
<tr>
<td>5.3</td>
<td>Simulation Results</td>
<td>95</td>
</tr>
<tr>
<td>5.4</td>
<td>Shopping on a Torus</td>
<td>98</td>
</tr>
<tr>
<td>6</td>
<td>Myopic Best-Reply with Mistakes</td>
<td>107</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>107</td>
</tr>
<tr>
<td>6.2</td>
<td>Noisy best-reply Dynamics</td>
<td>110</td>
</tr>
<tr>
<td>6.3</td>
<td>Payoff-dependent Mistakes</td>
<td>121</td>
</tr>
<tr>
<td>6.3.1</td>
<td>The Model</td>
<td>122</td>
</tr>
<tr>
<td>6.3.2</td>
<td>A &quot;local&quot; characterization of the process</td>
<td>124</td>
</tr>
<tr>
<td>6.3.3</td>
<td>Limit behaviour of the process</td>
<td>129</td>
</tr>
</tbody>
</table>
Aknowledgments

I wish to thank very many people I interacted with at different stages of the research leading to this dissertation, and shall do so in strictly chronological order.

In 1988 I wrote my (Italian) BA dissertation. It was Macroeconomics, and it involved policy coordination models. In one Chapter I constructed a model where I interpreted realignments among European currencies in terms of some form of Axelrod's model on the emergence of cooperation. I talked about "conventions" and "norms". That was pretty unusual, and different from all the formalizations that I diligently reviewed in another Chapter. I liked that simple model, and so did my supervisor (Prof. L. Spaventa). I had however to struggle a little bit to model my few country-players as very many and exactly identical. Also, although I was quite comfortable with the assumption that players were not necessarily payoff maximizers, I had some problems in moving away from the notion of learning, with which I had been brought up in a Bayesian Faculty of Statistics. After my MSc. in Economics, I wrote a research proposal. This time it was Microeconomics and it involved fundamental questions about learning and evolution in games. I started by looking at "naive" Bayesian players, whose priors sometimes converged to Nash-equilibrium. This was at Birkbeck College (University of London), where most of the people were working in Macroeconomics. Finding still unpublished papers in the field I was interested in was a struggle. My advisor at that time (Ben Lockwood), suggested me to talk to Tilman Börgers, who was a game theorist and had just joined University College London (University of London). I did so and I photocopied tons of papers. He also gave some lectures on the MPhil course, where he instilled in the audience the idea that equilibrium notions deprived of a learning process leading towards them was, at best, incomplete. In the meantime I started participating to the Theory Workshops
at University College London. Ken Binmore was organizing them. On the second workshop I attended, I was asked to present a paper by Larry Samuelson, and I had to read another paper, by Kandori, Mailath and Rob. These papers looked very different from those I had read before. I found them intriguing, though I was also a bit perplexed by an approach that was entirely new to me. Ken, in a brief informal conversation, convinced me that the idea of mutations could overcome the disturbing problems of priors that have zero measure in a learning model. I then traded off absolute continuity for "noise", and joined University College London. I found that a very stimulating environment and I profited from the interaction with virtually everybody over the four years I spent there, including eminent visitors such as Terence Gorman, in the office next door, with whom I had very long conversations. Giovanni Ponti, Raffaele Miniaci, Maria Saez-Marti, with whom I shared an office, many coffee breaks and a number of interests. Luca Anderlini, whom I had met during a visit to Cambridge as an undergraduate, gave a seminar at University College London. He had recently bought a Cellular Automata Machine that he showed me, while discussing about Turing machines. The interaction between the coloured screens and the paper by Kandori, Mailath and Rob I had just read, triggered further research. Two papers were the outcome of scribbling on the blackboard in the graduate students common room in Cambridge. A good part of the material of which in Chapters 4 and 5 is taken from there. I wish to thank him for the permission. I presented parts of the work in this dissertation in many occasions and in many places. At Summer in Tel Aviv 1993, among others, I found Eddie Dekel's comments terribly pertinent and very useful. In Vienna, at the I.I.A.S.A., I had long conversations with Peyton Young and a circle of people visiting that institution at the time. Later on in Mannheim, at the University of Pennsylvania, at CORE and in Bonn, I had the chance to meet many of the authors of the papers I was reading. I profited from very many informal conversations, that I remember vividly, but that it would take far too much space to summarize here. In a list that would, in principle, be far longer, I wish to name
Richard Boylan, Stephen Durlauf, John Geanakoplos, Bart Lipman and Ariel Rubinstein for interesting conversations we had in different places of the world at different times.

This dissertation was actually written during this last solar year. The material I had a year ago looked very different from the current version. For this value added, my gratitude uniquely goes to Tilman Börgers. Interacting with him has been very challenging, at time vibrant and constantly productive. To him, and to Ken Binmore, I owe my most sincere thanks.

Apart from myself, only one other person has, as yet, read every single word of this dissertation. No further words. Very many thanks.

It goes without saying that, whenever I cannot blame a co-author, I take the full responsibility of any remaining mistakes.

This thesis, along with many other things, is dedicated to my father:

A Penna Bianca,
con tutto l'amore del ricordo.
Introduction

Most economic applications of non-cooperative Game Theory build upon the concept of Nash-equilibrium. From a purely theoretical point of view, the underlying notion of rationality is recognized to be very demanding. On these grounds, an approach attracting increasing research activity analyzes the evolution of play. The idea underlying this line of research is that players, who are not assumed to be rational in any epistemologically well-defined sense, learn how to play the game by reacting adaptively to the environment around them. The focus is on the evolution of the sequence of play when a population of players interact repeatedly over time. In many cases of interest, convergence and stability properties of the dynamics have been shown to provide a robust foundation for the equilibrium notion.

The evolutionary paradigm hence relies on the characterization of the notion of Nash-equilibrium in terms of aggregate behaviour. Though not pursued explicitly, the idea dates back to Nash himself, who suggested a "mass-action" interpretation of the equilibrium notion (J. Nash, PhD Dissertation, Princeton University, 1950). Several approaches have been taken in the recent literature. Perhaps the best studied setting draws on the biological literature and analyzes deterministic selection dynamics. The underlying game is then played by couples of individuals, randomly chosen from an infinite population. Convergence properties of the dynamics are shown to constitute a robust rationale for Nash-equilibrium behaviour. Furthermore, asymptotic stability proves to characterize some equilibria as being "robust" to small perturbations, and as such more likely to be selected as the likely evolution of play.
However, many economic applications do not conform to the standard setting of the biological metaphor. Interactive contests often involve only a finite, and perhaps small, number of agents who take decisions on the basis of the limited amount of information they have as to the evolution of play in the whole population. As a result, the infinite population / random matching assumption is not always appropriate. Indeed, if agents' decisions were (stochastically) independent, based on a commonly shared "view of the world", then one would expect aggregate play to be close to the average play in the population. On the contrary, empirical evidence suggests that many phenomena of interest in economics show a persistently high cross-sectional variance.

Individuals may hold private "views of the world" for a number of reasons. If the population is very small, for example, the mere fact that a player excludes himself from the set of his potential opponents might alter the perception he has of the way the game is played in the aggregate. Besides, the random matching of individuals might not be uniform: for each player, interactions with some players in the population might be more likely than interaction with others. As it is often the case in real life situations, a player might be more likely to meet a subset of players as his opponents, and not interact at all with others. His decisions will then be conditional on his personal "view of the world". Strategy choices in the population would then show a certain amount of correlation that is ruled out by assumption in an entirely biological framework.

The intuitive appeal of these considerations is what motivates this dissertation. In particular, the questions we address can be informally summarized as follows.

The starting point of the analysis is that interaction patterns among players may differ from the standard biological metaphor. How can we then model a "population game" that accounts for that as a particular case? How can we define equilibria at the level of the aggregate population?

We would like to characterize equilibria at the population level in terms of equilibria
of the underlying game. If, and under which assumptions, can we interpret
the former as a "replicated" version of the latter? In other words, what are
the restrictions on the way interaction takes place at the population level that
allow for a the straight-forward identification of population equilibria in terms
of Nash-equilibria of the underlying game?

How can we model an evolutionary process of learning? Under what assumptions
would convergence be obtained? How could stochastic attractors be characterized?
How would they relate to the set of equilibria of the underlying game?

How robust would limit sets be to a modeled perturbation of the system? This
would account for mistakes that may occur while playing the game, as well as
for active experimentation on the part of players in an attempt to achieve a
better outcome. How can we model and explicitly analyze equilibrium selection
issues in this framework? How sensitive to the specification of the model would
the results be?

Description of the Thesis:

The thesis consists of 6 Chapters.

In Chapter 1 we provide a general formulation of the setting we analyze. In
particular, we specify a population game as being constituted by a set of players
(assumed to be finite throughout), an underlying normal form game (that we take to
be symmetric) for which we revise basic equilibrium concepts relevant to our analysis,
and an interaction pattern. The latter specifies the way players are matched in the
population to play one shot of the underlying game. The underlying idea (due to
Rosenthal (1979) and common to most of the literature in this field) is that each player
plays against a randomly chosen opponent. The probabilistic formulation we adopt
is fairly general and accounts for uniform population matching and local uniform matching as particular cases. On the lines of Mailath, Samuelson and Shaked (1994) and (1995), we relate the set of equilibria of the population game to the set of equilibria of the underlying game and show that the identification of the former in terms of the set of Nash-equilibria of the latter holds true in the limit, as the population becomes infinite. We prove that there exists a one-to-one correspondence between the set of equilibria of the population game and the set of symmetric correlated equilibria of the underlying game. Under uniform population matching and for the class of Coordination games, we show that if the population is finite, the equivalence between the set of strict equilibria of the population game and the set of strict Nash-equilibria of the underlying game relies on a further assumption on the payoffs of the underlying game. The assumption requires the relative advantage, in terms of expected payoffs, of one action over any other, be increasing in the probability that the opponent chooses the same action.

Chapter 2 provides an overview of the dynamic learning processes analyzed in later Chapters, as well as a general formulation of the stochastic dynamics. The first part complements this introduction by providing further motivations to this line of research. In essence, we take the view that a rationale for the equilibrium concepts of which in Chapter 1, may be provided by focusing on the process that leads to it. The learning processes we analyze are based on the idea that players' behaviour departs from the notion of rationality, based on Savage's (1954) axiomatization, traditionally used in Economics. We hold the view that a meaningful definition of non-rational behaviour can only be provided by focusing on the specific features of the economic or social setting being studied. Given that this work is entirely theoretical, the line we follow is to analyze many different non-rational learning processes. All the behavioural rules we analyze share as a common feature that individual behaviour is to be explained in terms of systematic, as well as random elements. We provide different interpretations of the latter in terms of inertia towards strategy changes
and/or mistakes (or trembles) affecting the decision process of players. We then postulate that interaction takes place repeatedly over time. Each of the remaining Chapters focuses on the learning process, at the population level, generated under each specific behavioural rule.

Chapter 3 focuses on myopic best-reply behaviour. The underlying idea is that players hold static expectations as to their opponents' choices and, whenever they are allowed to choose the action to be played, they aim at maximizing their expected payoff. We formulate general conditions on the population game, under which the process converges to an absorbing state for models of uniform population matching, as well as uniform local matching; we then interpret attractors of the dynamics in terms of equilibria of the underlying game. Though path-dependence applies throughout, for the class of Coordination games and under uniform population matching we are able to explicitly relate absorption probabilities to basic feature of the underlying game. To this aim, we provide a generalization of the notion of risk-dominance due to Harsanyi and Selten (1988), generic $m$-m Coordination games.

Chapter 4 analyzes majority rules of behaviour, where players are assumed to "imitate" the actions they observe in the population. We show that such rules are consistent with payoff maximizing behaviour if the game being played is a Coordination game for which all pure strategy Nash-equilibria are risk-equivalent. We show that cyclic limit behaviour is ruled out if players update their choices with inertia towards changes. We formalize this in terms of noise at the margin: players follow the rule with probability one, only if no change is prescribed, but may instead adopt any action with positive probability otherwise. Most of the convergence results for a local matching model rely on the possibility to choose, arbitrarily, the size of the neighbourhood (that is the same for all players and fixed over time). For the specific spatial arrangement of a Torus we show that such assumption may be relaxed.

Chapter 5 focuses on two simple processes of learning in a model of local uniform matching where players follow an aspiration level learning rule. The aspiration level
is defined by a deterministic payoff threshold, with which players compare the actual payoff obtained in the last round of interaction. We confine the attention to an underlying 2-2 Coordination game and we derive results from the findings of Chapters 3 and 4. We report on simulation results obtained with a Cellular Automaton Machine, that we use to characterize absorbing states of the process. Lastly, we interpret a version of the model as a model of local price search, showing that a highly decentralized interaction structure proves to support equilibrium price dispersion.

Chapter 6 analyzes models of myopic best-reply with mistakes. The first part of the Chapter reports on the approach common to most of the literature originated by Kandori, Mailath and Rob (1993) and Young (1993), analyzing learning processes for a normal form game, repeatedly played by a finite population of players. The dynamics are referred to as noisy best-reply, and formalize a perturbed version of the myopic best-reply of which in Chapter 3. The key feature of the model is that mistakes occur with a fixed, strictly positive probability, in a way that is independent of any payoff consideration, constant across players and over time. As a result, the learning process, at the individual, as well as at the population level is ergodic, meaning that its asymptotics do not depend on the initial condition. This motivates a further characterization of the process, as the "noise" becomes negligible, and allows to analyze equilibrium selection issues, by comparing the rates of convergence to different equilibria. We provide an example that casts doubt on the generality of the results when interaction takes place locally. In the second part of the Chapter we focus on a model where the probability with which mistakes occur does depend on payoff considerations and that mistakes that entail a high loss are relatively less likely to occur. For some classes of games we are able to characterize the limit distribution explicitly, as the Gibbs ensemble associated to a particular specification of the interaction potential, that we express by averaging expected payoffs across the population of players. Ergodicity properties of the stochastic process then motivate the study of equilibrium selection issues.
In the last part of the Introduction to most Chapters, there is a brief part that explicitly relates the work to the existing literature. The comments refer to other papers, often quoted in the text, but never described in detail, as well as to further personal conjectures. As a result, that part is completely obscure to whoever is not familiar with the references. It can be easily identified because indented and skipped without loss of continuity.

All opening quotations are taken from the prose of a great pre-Socratic philosopher, Heraclitus. I resisted the temptation to use the original language, and I acknowledge C.H. Kahn's translation in *The art and thought of Heraclitus*, Cambridge University Press (1987).
Chapter 1

Games and Equilibria

Let us not concur casually about the most important matters.

(XI)

1.1 Introduction

In this Chapter we provide a very general formalization of a population game. The framework is that of a collection of players randomly matched to play a normal form game. The key feature of the model, that distinguishes it from the specifications analyzed in later chapters, is that it is entirely static. The focus is hence on the characterization of the equilibria of the population game, and in particular on the relation between these, and the set of equilibria of the underlying game. The underlying question, that motivates the work and leads the logical exposition, is to what extent and under which assumptions the analysis of a population game can be carried out in terms of the study of the underlying game.

We take the underlying game to be symmetric and hence focus on a single population of players. The idea is that players are ex-ante exactly identical. An immediate implication is that the symmetry properties must necessarily be satisfied in equilibrium.
A key ingredient of a population game is the formalization of the way players are coupled. In this respect, consistently with the idea that in many settings of interest in economics, matching can follow highly arbitrary patterns, we focus on a very general framework. Matching is described by a probability distribution over unordered couples of players. As we point out standard specifications of matching, common to most of the literature in the field, are special cases.

The random matching device introduces a degree of correlation among players' equilibrium choices. On the lines of Mailath, Samuelson and Shaked (1994), we show that the latter can appropriately be accounted for by looking at the set of correlated equilibria of the underlying game. Our model differs, in that the focus is on a single population of players. The set of correlated equilibria is in general a wider set than the set of Nash equilibria. We show that, if matching is uniform, the set of symmetric Nash equilibria can be achieved, in the limit, by a sequence of population games for which the population size grows indefinitely.

We point out that the class of coordination games is particularly interesting with regard to the question we address, in that it provides an example of interacting situation where, looking at the set of Nash-equilibria of the underlying game does not in general exhaust the analysis of the equilibria of the population game.

The result of Theorem 1 formalizes the "additional complexity" of Footnote 3 in Mailath, Samuelson and Shaked (1994). Strictly speaking, the result of Theorem 1 in that paper does not apply to a symmetric underlying game (because the requirement of symmetry of the correlated equilibrium is invalidated in both part 1.1 and part 1.2 of that Theorem). In Mailath, Samuelson and Shaked (1995), a revised version of Section 2 of Mailath, Samuelson and Shaked (1994), the case of an underlying symmetric game is dealt with verbally on p. 5. The result of Theorem 1 in this Chapter provides an explicit formalization to that claim.
1.2 The general model

1.2.1 The players:

Let $\Omega \ni \omega$ be the set of players. On $\Omega$, we introduce the following assumption\(^1\):

**Assumption 1** $2 < \#\Omega < \infty$.

Assumption 1 is introduced in order for $\Omega$ to represent a finite population of players. We then assume that players are exactly identical and focus on the way interaction takes place. A *random matching* among players is a probability space over the set of unordered pairs of players. Let $\mu$ be a probability measure defined over the set of unordered pairs $P = \{\{\omega, s\} : \omega \in \Omega, s \in \Omega\}$. Then $\mu$ is a *random matching* if it satisfies the following assumption:

**Assumption 2** $\mu$ is such that $\forall \omega \in \Omega \exists s \in \Omega : \mu(\{\omega, s\}) > 0$.

Assumption 2 simply requires that each player in the population will, with some positive probability, take part in the interactive contest. The support of $\mu$ defines the *interaction pattern* among players. As far as the latter is concerned, a few specifications are worth being noted. A natural requirement is that players do not interact with themselves; clearly this amounts to further assume that $\mu(\{\omega, s\}) = 0$ for all $\omega = s$. Besides, given that couples of players are unordered, while a game requires the specification of the role, *row* vs. *column*, with which the interaction takes place, what is implicit in the contest is that players are equally likely to play in any of the two roles, or, in other words, that $\mu(\{\omega, s\}) = 2\mu((\omega, s))$, where the latter formalizes the convention that the first player, in the ordered pair, will play as the row. Lastly, note that assumption 2 does not rule out the possibility that the population of players is divided into a number of subsets of players: players might interact exclusively with players within the same group, or, alternatively, they might exclusively interact with

\(^1\)Here and throughout the whole work, the symbol $\#A$ stands for the cardinality of $A$. 

21
players who belong to a different subset. While the first case is an intrinsic replica of the original one, the second de facto reproduces one where different population of players interact with each other. If \( \mu \) splits the population in exactly two groups, say \( \Omega_I \) and \( \Omega_{II} \), we might for example assign roles to players according to the group they belong to\(^2\).

Any of the above specifications of the interaction pattern qualifies the support of \( \mu \). Let this be \( S \). Its cardinality can be found by simple counting argument and will clearly reflect the specific assumption under which interaction takes place. In general, a random matching is defined to be uniform, if the following holds:

**Definition 1** The random matching is uniform if, for all \( \{\omega, s\} \in S \), \( \mu(\{\omega, s\}) = \frac{1}{\#S} \).

Definition 1 allows for distinct assumptions about the matching technology. A model, commonly referred to in the literature as a random matching model, formalizes the idea that each player is equally likely to be matched with any other player in the population. The latter is obtained in this context by assuming that \( \mu \) has full support, i.e. \( S = P \). We will refer to this case as uniform population matching, and we will denote it by \( \mu_U \). If, instead, players interact only with a subset of players in the population, i.e. \( S \subseteq P \), then Definition 1 is to be read in terms of what is customarily referred to as a model of locally uniform matching. The idea underlying this kind of models is that the set of potential opponents for player \( \omega \) depends on \( \omega \) itself, and player \( \omega \) is equally likely to meet any player within that set. A way to distinguish player \( \omega \) from player \( s \) in the population is to provide each player with an "identity" in terms of a specific location on an appropriately characterized space \( \Omega \). We will explicitly formalize models of locally uniform interaction, in Chapter 4, 5 and 6. For the purposes of this Chapter, we only point out that a uniform random

\(^2\)A convenient, though not essential for the purposes of this Chapter, representation of \( \Omega \) is in terms of a graph, \( G(\Omega, E) \), where \( E \) is the set of edges connecting the elements of \( \Omega \) for which \( \mu(\{\omega, s\}) > 0 \). Whenever we refer to the graph representation, as we will extensively do in Chapter (4), we might characterize the interaction pattern in more detail, by introducing further assumptions on \( G \). In this framework, a two population model can be derived by assuming that \( G \) is bipartite.
matching model that rules out self-interaction can be thought of as an extreme version of a locally interactive model, by the mere fact that the set of opponents that $\omega$ faces excludes $\omega$ himself.

Each player $\omega$ chooses an action among those available to him, to play an underlying game.

1.2.2 The game:

Each player $\omega$ can choose any action in the finite set $A$, of cardinality $m$, to play a normal form game $G$. We take $G$ to be symmetric; as such $G$ is described by a single $m \times m$ payoff matrix:

Assumption 3 $G$ is a symmetric normal form $m$-$m$ game:

$$G \equiv [\pi_{ij}] \quad \pi_{ij} \in \mathbb{R} \quad \forall i, j = 1, \ldots, m$$

This simply means that, if player $\omega$ chooses action $a(\omega) = i$ and player $s$ chooses action $a(s) = j$, then $\pi_{ij}$ is $\omega$'s payoff. In general, $\omega$'s payoff will be denoted as $\pi(a(\omega), a(s))$.

Being finite, $G$ admits at least one Nash-equilibrium in pure or in mixed strategies. In general, let $N(G)$ be the set of Nash-equilibria of $G$ and, furthermore, let $\Psi(G)$, as introduced in Aumann (1974), be the set of correlated equilibria of $G$. We focus on symmetric equilibria, as in the following definitions:

Definition 2 A symmetric Nash-equilibrium of $G$ is a probability distribution $p : A \rightarrow [0, 1]$ such that, for all actions $i \in A$ for which $p_i > 0$, the following holds:

$$\sum_{j=1}^{m} p_j (\pi_{ij} - \pi_{il}) \geq 0 \quad \forall l \neq i$$

$\tilde{N}(G) \subseteq N(G)$ defines the set of symmetric Nash-equilibria of $G$. 23
\( \bar{N}^*(G) \subseteq \bar{N}(G) \) defines the set of strict\(^3\) symmetric Nash-equilibria of \( G \).

**Definition 3** A symmetric correlated equilibrium of \( G \) is a probability distribution, \( \psi : A \times A \to [0,1] \) such that for all \( i \neq j \) \( \psi_{ij} = \psi_{ji} \), and for all actions \( i \in A \) for which \( \sum_j \psi_{ij} > 0 \), the following holds:

\[
\sum_{j=1}^{m} \psi_{ij} (\pi_{ij} - \pi_{ji}) \geq 0 \quad \forall i \neq j
\]

\( \tilde{\Psi}(G) \subseteq \Psi(G) \) defines the set of symmetric correlated equilibria of \( G^4 \).

\( \tilde{\Psi}^*(G) \subseteq \tilde{\Psi}(G) \) defines the set of degenerate\(^5\) symmetric correlated equilibria of \( G \).

We now move on to the population game.

### 1.2.3 The population game:

A population game is simply defined by the population itself, the way couples of players are formed and the underlying game that players play once matched.

**Definition 4** Given a population of players \( \Omega \), a random matching technology \( \mu \), and an underlying game \( G \), if \( \Omega, \mu \) and \( G \) satisfy Assumptions (1), (2) and (3) respectively, then the triple \( \Gamma = (\Omega, G, \mu) \) defines the population game.

In the remaining part of this Chapter, we analyze an entirely static model and we focus on the characterization of equilibria. In the framework of later Chapters, these are to be interpreted in terms of configuration of stable states that a dynamic process might reach asymptotically.

---

\(^3\)A Nash-equilibrium is strict if Definition (2) holds with the strict inequality sign.

\(^4\)It is customary to describe correlated equilibria by means of the metaphor of a referee recommending an action to each player. Players do not observe what action the referee recommends to other players. The equilibrium condition then requires that, given the recommendation, such actions be optimal for each player. Focusing on symmetric correlated equilibria amounts to postulate that such recommendations cannot be made conditional on the role that each player will have in the interaction.

\(^5\)A symmetric correlated equilibrium is degenerate if there exists an \( i \) such that \( \psi_{ii} = 1 \).
A configuration of actions in the population is \( \theta \), an element of the space \( \Theta = A^{2^n} \). Let \( a_\theta(\omega) \) be the action adopted by player \( \omega \) in the configuration of play \( \theta \). Then an equilibrium of the population game \( \Gamma = (\Omega, G, \mu) \) is defined as follows:

**Definition 5** \( \theta \) is an equilibrium of \( \Gamma \) if for each player \( \omega \in \Omega \) and for all actions \( \tilde{a}(\omega) \in A \), \( \tilde{a}(\omega) \neq a(\omega) \) the following holds:

\[
\sum_{\{\omega,s\}} \pi(a_\theta(\omega), a_\theta(s)) \mu(\{\omega,s\}) \geq \sum_{\{\omega,s\}} \pi(\tilde{a}_\theta(\omega), a_\theta(s)) \mu(\{\omega,s\})
\]

\( \Theta(\Gamma) \) defines the set of equilibria of \( \Gamma \).

Note that Definition 5 is independent of the role that each player has, while playing \( G \). Furthermore note that, given the degree of arbitrariness implicit in the definition of \( \mu \), it is perhaps more intuitive to re-state the above definition in terms of conditional probabilities:

\[
\mu(s \mid \omega) = \frac{\mu(\{\omega,s\})}{\sum_s \mu(\{\omega,s\})}
\]

to emphasize the fact that what is relevant to player \( \omega \) are the actions chosen by his potential opponents. We will often refer to \( \mu(s \mid \omega) \) as player \( \omega \)'s "view of the world". In general, players might hold different views of the world. Even if matching is uniform, the mere fact that a player excludes himself from the set of potential opponents introduces a bias in the perception of how the game is played at the population level\(^6\).

Each \( \theta \in \Theta \) induces a partition over the set of players, \( \Omega \), and over the set of couples of players, \( S \), according to the actions that are chosen. The restrictions of \( \mu \) over these partitions are defined as \( \mu^\theta_A \) and \( \mu^\theta_{A \times A} \) respectively. Intuitively \( \mu^\theta_A \) and

\(^6\)It is clear that the only case where such correlation between players' views of the world fades away is if \( \mu \) is specified as uniform over \( \Omega \times \Omega \), i.e. if we take the support of \( \mu \) to include all ordered couples of players, allowing for repetition.
describe the frequencies with which an outside observer would see strategies and
strategy pairs represented in the population of players, randomly matched according
to µ, when each player is playing as in θ. By the definition of µ, the support of
µAxA contains only m(m + 1)/2 elements (because the "order" with which different
actions are chosen cannot be accounted for). Instead, we find it convenient to define
a probability distribution over the m² elements of A × A. To this aim, given θ, let
us denote the elements of the partition as: {i, i}θ = \{\{ω, s\} : aθ(ω) = aθ(s) = i\}
and, for j ≠ i, {i, j}θ = \{\{ω, s\} : aθ(ω) = j, aθ(s) = i\}. Then we define µAxA as
follows:

\[ µ_{A \times A} = \begin{cases} 
    µ(\{i, i\} θ) & \text{if } i = j \\
    \frac{1}{2} µ(\{i, j\} θ) & \text{if } i ≠ j
\end{cases} \]

The above definition relies on the implicit assumption that each player is equally
likely to play the game G as the row player, or as the column player. Clearly, in
aggregate, the frequencies with which actions are to be adopted in the population
correspond to the marginal distribution of µAxA: the share of action i in µA is in fact
given by:

\[ µ(\{i\} θ = \{ω ∈ Ω : aθ(ω) = i\}) = µ(\{i, i\} θ) + \sum_{j ≠ i} \frac{1}{2} µ(\{i, j\} θ) \]

We will refer to \{µAθ : θ ∈ Θ(Γ)\} = M(Γ) and to \{µAxA : Θ(Γ)\} = M(Γ) as the
equilibrium frequencies of the population game Γ. The purpose of the Section that
follows is to relate these sets with the set of equilibria of the underlying game.

1.3 Equilibria of the game and equilibria of the population game.

We now focus on the relation between the set of equilibria of Γ = (Ω, G, µ) and
the set of equilibria of G. The remark that follows shows that each equilibrium of
the population game induces, in terms of frequencies over ordered pairs of actions, a 
correlated equilibrium of the underlying game. The proof reproduces the one provided in Theorem 1 of Mailath, Samuelson and Shaked (1994). The model we analyze differs 
from the latter in that we deal with a single population of players\(^7\).

**Remark 1** \(\mathcal{M}(\Gamma) = \hat{\Psi}(G)\).

**Proof.**

\(\mathcal{M}(\Gamma) \subseteq \hat{\Psi}(G)\) : Assume \(\theta \in \Theta(\Gamma)\). Then by Definition 5, for each player \(w\) adopting action \(i\) in equilibrium \(\theta\) the following must hold:

\[
\sum_{j=1}^{m} \sum_{s \in \Omega \text{ s.t. } a_{\theta}(s) = j} \mu(\{w, s\})(\pi_{ij} - \pi_{lj}) \geq 0 \quad \forall l \neq i \quad (1.1)
\]

Summing (1.1) over all players \(\omega\) s.t. \(a_{\theta}(\omega) = i\):

\[
\sum_{j=1}^{m} \sum_{\omega \in \Omega \text{ s.t. } a_{\theta}(\omega) = i} \sum_{s \in \Omega \text{ s.t. } a_{\theta}(s) = j} \mu(\{\omega, s\})(\pi_{ij} - \pi_{lj}) \geq 0 \quad \forall l \neq i \quad (1.2)
\]

or:

\[
\sum_{\omega \in \Omega \text{ s.t. } a_{\theta}(\omega) = i} \sum_{s \in \Omega \text{ s.t. } a_{\theta}(s) = i} \mu(\{\omega, s\})(\pi_{ii} - \pi_{ii}) + \sum_{\omega \in \Omega \text{ s.t. } a_{\theta}(\omega) = i} \sum_{s \in \Omega \text{ s.t. } a_{\theta}(s) = j} \mu(\{\omega, s\})(\pi_{ij} - \pi_{lj}) \geq 0 \quad \forall l \neq i \quad (1.3)
\]

which, for \(l \neq j \neq i\), in turn yields:

\[
\sum_{\{\omega, s\} \in S \text{ s.t. } a_{\theta}(\omega) = a_{\theta}(s) = i} \mu(\{\omega, s\})(\pi_{ii} - \pi_{li}) + \frac{1}{2} \sum_{\{\omega, s\} \in S \text{ s.t. } a_{\theta}(\omega) = i a_{\theta}(s) = j} \mu(\{\omega, s\})(\pi_{ij} - \pi_{ij}) \geq 0 \quad (1.3)
\]

\(^7\)Mailath, Samuelson and Shaked (1994)'s model is fully recovered if we introduce as a further assumption on \(\mu\), the requirement that the population is partitionned into two groups, and roles are assigned to players according to the group they belong to.
The reasoning can be extended to all actions adopted in equilibrium $\theta$, thus reproducing the equilibrium frequencies of which in $\mathcal{M}(\Gamma)$. By Definition 3, inequalities as in (1.3) reproduce a correlated equilibrium of $G$ once we take:

$$\psi_i = \sum_{\{\omega,s\} \quad s.t.\text{ } a_\theta(\omega) = a_\theta(s) = i} \mu(\{\omega, s\})$$

and

$$\psi_{ij} = \frac{1}{2} \sum_{\{\omega,s\} \quad s.t.\text{ } a_\theta(\omega) = i \text{ and } a_\theta(s) = j} \mu(\{\omega, s\})$$

The symmetry requirement that stems from the definition of $\mu$ implies that $\psi_{ij} = \psi_{ji}$.

$\mathcal{M}(\Gamma) \supseteq \hat{\Phi}(G)$: Given a symmetric correlated equilibrium of $G$, we need to show that it is possible to define a probability distribution $\mu$ over the set of unordered pairs of elements of a finite set $\Omega$, such that Assumption 2 is satisfied. The latter only requires that for all $\omega \in \Omega$, $\mu(\omega) > 0$. If $m^* \leq m$ is the number of actions $i \in A$ for which $\sum_j \psi_{ij} > 0$, we can always take a population of $2m^*$ players divided into two groups of equal cardinality, assigning one available action $i$ to each member of each group and defining $\mu$ so that:

$$\mu(\{i,j\}) = \begin{cases} \psi_i & \text{if } j = i \\ \psi_{ij} + \psi_{ji} & \text{if } j \neq i \end{cases}$$

The notion of correlated equilibrium captures the amount of correlation between strategy choices that may obtain in an equilibrium configuration. The intuition is that player $\omega$'s equilibrium choices depend on his "view of the world" which might not be the same for different players. The combination of large population kinds of
arguments and uniform matching technology, allows to approximately disregard such correlation and motivates the interpretation of equilibria of the population game in terms of replicated versions of the symmetric Nash-equilibria of the underlying game, as shown in the Remark that follows.

Remark 2 Consider a sequence $(\Gamma^n)_{n \in \mathbb{N}}$ of population games $\Gamma^n = (\Omega^n, G^n, \mu^n)$ such that for each $n \in \mathbb{N}$:

(i) $G^n = G$,

(ii) $\mu^n = \mu^n_0$,

(iii) $\lim_{n \to \infty} \#\Omega^n = \infty$.

For each $n \in \mathbb{N}$ let $\mu^n_{A \times A} \in \mathcal{M}(\Gamma^n)$. Assume that $\lim_{n \to \infty} \mu^n_{A \times A} = \mu_{A \times A}^\infty$.

Then $\mu_{A \times A}^\infty \in \hat{\Psi}(G)$ and $\mu_{A \times A}^\infty \in \bar{N}(G)$.

Proof. Since the set of correlated equilibria is closed, $\mu_{A \times A}^\infty$ is a correlated equilibrium. It therefore suffices to show that $\mu_{A \times A}^\infty$ is the product measure of the marginal distribution $\mu_{A}^\infty$. To lighten notation, let $n \equiv \#\Omega^n$ and, for a given equilibrium configuration of play $\theta^n = \theta(\Gamma^n)$, $n_i = \# \{ \omega \in \Omega^n : a_{\theta^n}(\omega) = i \}$, i.e. the number of players that in an equilibrium of the population game $\Gamma^n$ are adopting action $i$. By definition:

\[ i \in A \Rightarrow \mu^n_{A \times A}(\{i,i\}) = \frac{n_i n_i - 1}{n (n - 1)} \]

Hence:

\[ \lim_{n \to \infty} \mu^n_{A \times A}(\{i,i\}) = \left( \lim_{n \to \infty} \frac{n_i}{n} \right)^2 \]

\[ ^8\text{We use } \mathbb{N} \text{ to denote the set of Natural numbers.} \]

29
Similarly:

\[ i, j \in A, i \neq j \Rightarrow \mu_{A \times A}^{n}(\{i, j\}) = \frac{n_i \cdot n_j}{n \cdot n - 1} \]

Hence:

\[ \lim_{n \to \infty} \mu_{A \times A}^{n}(\{i, j\}) = \left( \lim_{n \to \infty} \frac{n_i}{n} \right) \left( \lim_{n \to \infty} \frac{n_j}{n} \right) \]

thus proving the assertion.

If the population is large and matching is uniform, the frequencies with which different actions are adopted in the population reproduce the probability coefficients that yield any of the symmetric Nash-equilibria of the underlying game. The equivalence between the set of equilibria of \( G \) and the set of equilibria of the population game is, in this sense straight-forward, in that the latter can be thought of as "replicated" versions of the former. By looking at the equilibria of \( G \), we would therefore implicitly consider all equilibria of the associated population game, \( \Gamma \). On the other hand, in general if the population is finite and in particular if matching is not uniform, a consistent interpretation of the population game \( \Gamma \) must focus on the (wider) set of symmetric correlated equilibria of \( G \), where the correlation device is given by the matching technology, as well as by the equilibrium condition.

Analyzing the set of equilibria of the underlying game is not often an easy task. However, for some classes of games the characterization is particularly simple. For reasons that will become clearer in later Chapters, we will find it useful to restrict the attention to a subset of equilibria. In particular, while by definition a Nash-equilibrium is contained in its set of best-replies, we introduce as a further requirement that it also contains the latter set. This amounts to say that we focus on the degenerate probability distribution that defines a strict Nash-equilibrium. Strict equilibria have the specific property that unilateral deviations are necessarily costly, and as such are necessarily pure-strategy equilibria. We extend the refinement to the
set of correlated equilibria by focusing on the subset for which the defining probability distribution is degenerate. For essentially the same reason why, for a symmetric game, strict equilibria are not necessarily to be found on the main diagonal of the payoff matrix, degenerate correlated equilibria may not be symmetric. As it is clear from the Definitions we gave in Section 1.2, if we focus on pure strategy equilibria, the equilibrium definitions can be re-stated in terms of explicit conditions on the payoffs of the game. The class of games for which we explicitly pursue this line, is the class of Coordination games, to which we devote the next Section.

1.4 Coordination Games

A symmetric coordination game, $C$, satisfies the following Definition:

**Definition 6** $C_m$ is a coordination game if for all $i \in A$, $\pi_{ii} > \pi_{ji}$ $\forall j \neq i$.

Definition 6 states that payoff is maximized when both players choose the same action. As a result, any probability distribution that assigns probability one to only one action constitutes a strict Nash-equilibrium of $C_m$. For exactly the same reason, any probability distribution that assigns probability one to $(i,i)$ is a symmetric correlated equilibrium of $C_m$. While in general $\tilde{N}(C_m) \subseteq N(C_m)$ in the same way as $\tilde{\Psi}(C_m) \subseteq \Psi(C_m)$, by focusing on pure strategy equilibria the same relation holds with the equality sign, as we explicitly state in the next Remark:

**Remark 3** For a coordination game $C_m$

\[ \tilde{N}^*(C_m) = N^*(C_m), \quad \tilde{\Psi}^*(C_m) = \Psi^*(C_m). \]

We now turn to the associated population game, $\Gamma = (\Omega, C, \mu)$. As far as the latter is concerned, the results of Remark 1 and 2 are of straight-forward application. If the matching technology is uniform, and the population size is infinite, then the only
equilibria of the associated population game are those configurations of play where the frequencies with which available actions are chosen in the population reproduce any of the Nash-equilibria of the underlying game. However, if the population is finite and/or matching is local, the relevant set of equilibria to be looked at is that of the symmetric correlated equilibria of the underlying game, and these might not reproduce any Nash-equilibrium of the underlying game. For example, in a uniform population matching model for a 2-2 coordination game we know that the only equilibria involve all players adopting the same action. This is no longer the case if the matching technology is arbitrarily specified, in that players might hold, in equilibrium, different "views of the world" and on the basis of the latter, choose different actions. As a result, both action can coexist in equilibrium.

We conclude this Section with a general remark on the latter point, where we address the following question. We know from Remark 2 that, if matching is uniform, the set of equilibria of the population game can be identified, in the limit, with the set of symmetric Nash equilibria of the underlying game. Clearly the remark holds also if we require players to be adopting a strict best-reply, in which case in the limit we would only observe frequencies that reproduce degenerate equilibria. Given that the identification of the latter set is particularly simple for an underlying coordination game, we wonder whether by introducing further assumptions on the payoff structure, we can obtain the same result for a finite population.

To clarify, consider $\Gamma = (\Omega, C_2, \mu_U)$. Assume in an equilibrium of $\Gamma$ there is at least player $\omega$ adopting the first action. Then it must be that, for $\omega$:

\[(\pi_{11} - \pi_{21})\frac{\#\Omega_1 - 1}{\#\Omega - 1} + (\pi_{12} - \pi_{22})\frac{\#\Omega - \#\Omega_1}{\#\Omega - 1} \geq 0\]

\[^{9}\text{In Chapters 3 and 4 we will "visualize" a few examples.}\]
Can there be a player $s$ adopting, in equilibrium action 2? The answer is no: $s$ would in fact adopt action 2 in equilibrium only if:

$$(\pi_{11} - \pi_{21})\frac{\#\omega_1}{\#\omega - 1} + (\pi_{12} - \pi_{22})\frac{\#\omega - \#\omega_1 - 1}{\#\omega - 1} \leq 0$$

but this cannot be the case, because by adding $\frac{1}{\#\omega - 1}(\pi_{11} - \pi_{21} - \pi_{12} + \pi_{22}) > 0$ under Assumption 6) to the first equilibrium condition we contradict the second. Hence Assumption 6 is necessary and sufficient to rule out equilibria where both action are represented in the population. This is no longer the case for a generic $m$-by-$m$ coordination game, as shown in the Remark that follows.

**Remark 4** Given $\Gamma = (\Omega, C_m, \mu_U)$, $\mathcal{M}(\Gamma) \supseteq \hat{\Psi}^s(C_m)$ and $\mathcal{M}(\Gamma) \supseteq \hat{N}^s(C_m)$. If for all $i \neq j \neq l$, $(\pi_{ii} - \pi_{il}) > (\pi_{ij} - \pi_{ij})$, then $\mathcal{M}(\Gamma) = \hat{\Psi}^s(C_m)$ and $\mathcal{M}(\Gamma) = \hat{N}^s(C_m)$.

**Proof.** The $\supseteq$ is obvious. For the $\subseteq$ in the second part of the statement, it is sufficient to prove that, given an equilibrium $\theta$, if $a_\theta(\omega) = i$, then $a_\theta(s) = i$ for all $s \in \Omega$. If $\omega$ chooses $i$, then:

$$(\pi_{ii} - \pi_{il})\frac{\#\omega_4 - 1}{\#\omega - 1} + \sum_{j \neq i} \frac{\#\omega_j}{\#\omega - 1} (\pi_{ij} - \pi_{lj}) \geq 0 \ \forall l \neq i$$

By adding $\frac{1}{\#\omega - 1}(\pi_{ii} - \pi_{ii} - \pi_{ij} + \pi_{ij}) > 0$ we get, for all actions $j \neq i$:

$$(\pi_{jj} - \pi_{ij})\frac{\#\omega_j - 1}{\#\omega - 1} + \sum_{j \neq i} \frac{\#\omega_4}{\#\omega - 1} (\pi_{ji} - \pi_{il}) < 0 \ \forall l \neq j$$

which rules out optimality for any action $j \neq i$. $lacksquare$

The above Remark shows that, if the population is finite, a further assumption on the underlying payoffs (besides the one that defines a coordination game) allows to characterize equilibria of the population game exclusively in terms of strict equilibria.
of the underlying game. The condition requires that the comparative advantage of action \( i \) over action \( l \) be increasing in the probability with which a player's opponent adopts action \( i \). To see this, let \( q \) be any probability distribution over \( A \) such that 
\[ \sum_{j} q_j (\pi_{ij} - \pi_{lj}) \geq 0 \quad \forall l \neq i. \]
The latter function quantifies the relative advantage of action \( i \) over any action \( l \neq i \). Its partial derivative with respect to \( q_i \) is given by 
\[ (\pi_{ii} - \pi_{ii}) - (\pi_{ij} - \pi_{ij}) \quad \text{for any } j \neq i \]
and is strictly increasing in \( q_i \) if the above condition holds. In other words, if action \( i \) is the best-reply to \( q \), then this is also true for any \( \tilde{q} \) such that \( \tilde{q}_i > q_i \) and \( \tilde{q}_j \leq q_j \quad \forall j \neq i \). If the population is finite and matching is uniform, one player who adopts action \( i \) as his best reply faces a proportion of \( \frac{n_i - 1}{n - 1} \) other \( i \)-players. As it is clear, any other player will observe at least the same proportion of \( i \).

In the Chapter that follows we provide a general introduction to the learning processes that constitute the focus of this work.
Chapter 2

The Evolution of Play: an Overview

There is a certain order and fixed time for the change of the cosmos
in accordance with some fated necessity.

(CXIV)

This Chapter contains two Sections. In the first we motivate the analysis of the
line taken in the Chapters that follow; in the second we introduce the main features
of the models that will be analyzed.

2.1 Overview

Equilibrium concepts of the kind formalized in Chapter 1 are at the center of non-
cooperative Game Theory. In particular, Nash-equilibrium is commonly used in most
economic applications. Its logical foundations have been questioned closely in the
recent literature\textsuperscript{1}. Following a taxonomy used by K. Binmore in a number of works\textsuperscript{2},

\textsuperscript{1}Among other, and in strictly chronological order, we mention the works of Tan and Werlang
\textsuperscript{2}For example in K. Binmore (1987a) and (1987b).
two main approaches can be identified.

Within the so-called *eductive* approach, Nash Equilibrium is justified on the basis of pure introspection: deductive reasoning on the part of players produces the knowledge of equilibrium strategies that are unerringly identified and played. Refinements of the Nash Equilibrium concept provide an attempt to identify a unique plausible outcome in cases of multiplicity. The troubling feature of this approach is that agents are credited with an overabundant capacity for logical reasoning. Moreover, the assumption of common priors demands a high degree of coordination among player.

An alternative approach analyzes the *evolution* of play itself: instead of inferring players' behavior from an equilibrium notion, some plausible rules of behavior on the part of agents are postulated and the interest is then focused on the evolution of the sequence of play when interaction is repeated over time. Does the process converge? If so, does it converge to a Nash equilibrium? Can equilibrium selection issues be addressed? The idea underlying these models is that players learn how to play the game by reacting adaptively to the environment around them. The approach based on learning and evolution can potentially overcome the difficulties we mentioned above, in that it does not place unreasonable demands on players and it focuses on the dynamics of play itself, rather than on that of belief revision. The obvious drawback is the high degree of subjectivity that ensues from the rejection of the unbounded rationality/common prior assumption.

The approach we take in this work falls into the latter category, in that we analyze evolutionary processes of learning. Common features of the models that analyze learning in normal form games where the population is finite, can briefly be outlined as follows. A population of players are repeatedly and randomly matched to play

---

3 The literature in this field has been growing very fast in recent years. *Games and Economic Behavior* published a double Special Issue on Adaptive Dynamics in 1993 (Vol. 5, Nos. 3 and 4). Perhaps the most cited works appeared in the 61-th issue of *Econometrica*, authored by Kandori, Mailath and Rob and Young respectively. We shall explicitly refer to these works in Chapter 6.
a one-shot normal form game. At each round each player decides how to update its strategy choice for the next round on the basis of the information gathered about the past history of play. Such information generally consists of a sample of outcomes of past play or, typically, of some statistics of how the last round was played. Players are assumed to adopt some boundedly rational "rules of thumb", according to which they decide what to do next. The dynamics of the population as a whole is entirely determined by the intertemporal link stemming from the modelled rules. The interest is then focused on the long run properties of such system, and, if convergence is obtained, on the characterization of steady states of the process.

The full rationality paradigm is therefore abandoned in favor of behavioral rules that are myopic, in the sense of not being forward looking, and show some sort of inertia, in that not all agents update their strategy choices at the same time. The myopia assumption refers to the fact that agents act in a way that maximizes their expected payoff in the current period of play. The assumption is essential as the purpose of the analysis is to provide some rationale for equilibrium, and such possibility would be ruled out almost by definition if dynamically optimal behavior was postulated. The inertia assumption reflects the idea that changing strategy may be costly, or that the knowledge that agents have about payoffs or about the strategic behavior of their opponents is at best uncertain. The myopia assumption introduces an explicit path-dependence over strategy choices that governs the dynamics of the system as a whole and the inertia assumption further specifies the population dynamics.

In most of the literature to date, the next step in the analysis is to investigate what happens if the system is constantly hit by small random shocks that are independent across players and over time and have full support in the strategy space. Such random perturbations are meant to model experimentation and/or mistakes in strategy choices on the part of players, and are the precise analog of the basic notion of mutations, in biological evolutionary theory. In many cases such a formalization allows to derive an invariant distribution, that is a distribution over all possible states
towards which the stochastic process converges irrespective of the initial condition. Ergodicity properties then motivate the analysis of equilibrium selection issues in terms of limit properties of this distribution, for the modeled perturbation becoming infinitesimally small.

Following this line of argument, this Chapter is meant to introduce the different learning rules we analyze in this work.

2.2 The Evolution of play

Given a population game $\Gamma = (\Omega, G, \mu)$, we assume that the matching technology, as well as the underlying game, are time independent. Both are known to each player.

Time is discrete. In most of this thesis, we shall assume that, at each time $t$ only one couple of players actually play the game (though, by assumption, all players have a strictly positive probability to take part in the interaction). The latter will involve player $\omega$ and player $s$ with probability $\mu(\{\omega, s\})$. Before the random draw takes place, each player has chosen an action from the set $A$, by appealing to a rule. Within the same model, all players appeal to the same rule. For each player, a rule is simply a mapping from "what he knows" to "what he does". We shall describe different rules below. At the beginning of time, nature has assigned an action to each player. The first interaction takes place at time zero. Between time zero and time one some players receive an updating opportunity; the latter consists of the possibility to change the original action according to the rule. We shall specify different ways updating opportunities are allocated to players. Once all updating opportunities are used, time one's interaction takes place. As time rolls by, the above story is repeated with exactly the same timing.

We shall formalize two different ways updating opportunities are allocated to players:

- The first assumes that all players get a possibility to update their strategy
choices simultaneously at each discrete time steps. It will be referred to as synchronous (or parallel) dynamics.

- The second is the asynchronous (or sequential) dynamics, in which players update their strategy choices one-at-a-time, in some random order.

We devote the Section to the specification of the rules players follow when they get the possibility to choose an action.

2.2.1 Behavioral Rules

We assume that, when taking decisions, all players follow the same behavioral rule; the latter is simply a mapping from the information a player has into the set of actions. Whenever the rule produces multiple prescriptions, it will take the form of a probability distribution over the set of actions. At the beginning of time period $t$, all players are informed about the configuration of actions in the population at time $t - 1$ and/or (according to the specific rule) about the outcome of play.

In principle, a clear distinction can be made between behavioral rules that are entirely deterministic and behavioural rules that are specified in probabilistic terms. For expository purposes, we will list them in an increasing order of the weight of the stochastic component. Given that a detailed specification of each single rule will be given later, the list that follows, is only partly rigorous. Let $\Omega_\omega$ be the set of player $\omega$'s potential opponents, and $a(\omega)$ the action adopted by player $\omega$. We shall refer to $\mu_A(\omega)$ as the vector of frequencies with which different actions are chosen in $\Omega_\omega$ (a generic entry being $\mu_j(\omega)$ for $j \in A$); to $\pi_\omega(\omega)$ as $\omega$'s expected payoff; to $\pi(\omega)$ as $\omega$'s actual payoff. Actual payoffs come from the payoff matrix $[\pi_{ij}]$ of the underlying game $G_m$ that has action space $A$. Any time reference will be specified by a time subscript or superscript.

If we abstract from ties, the following are entirely deterministic rules:
• The first deterministic behavioral rule we analyze is known as the *myopic best-reply*, and formalizes the idea that players hold static expectations about their opponents’ behavior and, whenever they have the opportunity to do so, they choose the action that maximizes their expected payoff. As a result, if it is player ω’s updating turn, he will choose:

\[ a_ω(ω) \in \operatorname{Arg\max}_A \{π_ω(ω)\} \]

where the expected payoff to ω from action \( i \) is \( π^i_ω(ω) = \sum_{j \in A} π_{ij} μ^{i-1}_j(ω) \).

• The second rule we analyze is based on the idea that players refrain from performing any explicit optimizing procedure and simply *imitate* the actions that other players have adopted in the previous round. One specification we will analyze in detail is given in terms of *majority rules*, according to which, if it is player ω’s updating turn, he will choose:

\[ a_ω(ω) \in \operatorname{Arg\max}_A μ^{i-1}_A(ω) \]

The underlying idea is that behavior encountered in reality often are not the outcome of any analytical calculation, but instead motivated by some "herd" phenomena. It should be stressed that a peculiar feature of the models we analyze is that players imitate actions that have been adopted by other players, *irrespective* of the consideration of the payoff those action have achieved. This distinguishes the model from other specifications of imitation processes in an interactive contest, as well as from the literature on "herding" models or (purely) social learning, where learning is driven by pure informational externalities, in that an agent’s payoff is unaffected by the action of other agents.

• The third rule we analyze is meant to formalize the idea that players’ behavior
is partly driven by the reward obtained in the last round of interaction. Along this line, we will introduce further assumptions on the underlying model to describe a situation where all players play at least one shot of the underlying game in each period. In particular, we postulate that players have an aspiration level in terms of payoff, i.e. a minimum level above which they would consider themselves to be satisfied and they would stick to the action previously adopted.

\[ a_t(\omega) = a_{t-1}(\omega) \text{ if } \pi^{t-1}(\omega) > \bar{\pi} \]

where \( \bar{\pi} \) is the (for the time being unmodelled) aspiration level. We shall assume that the latter is a simple function of the payoffs of the underlying game; as such it will be constant across players and over time. In Chapter 5 we shall specify in different ways what happens if the reward falls below that threshold.

Note that though the above rules are essentially deterministic, at the level of the whole population the dynamics they generate is nevertheless stochastic. This is clearly so if updating is asynchronous (as in the specification we provided), and unless tie-breaking rules are specifically postulated to deal with multiple prescriptions, also in the synchronous dynamics. We define a generic rule of the above type as noiseless and to the aggregate process it generates as noiseless dynamics. We now move on to the stochastic rules we will analyze; each of the latter describes players' behaviour in probabilistic terms. We will refer to the stochastic process generated under a stochastic rule as a noisy dynamics. For the sake of exposition, a rule will be referred to as the mapping \( \mathcal{R} : \Theta \rightarrow A \) (one-to-many for stochastic rules): \( \mathcal{R}(\theta) \) is then what the rule prescribes if the configuration of play in the population is \( \theta \). The associated dynamics are denoted as \( \varphi^R \).

- The first stochastic behavioural rule we study is meant to formalize some inertia towards changes that might affect players choice. In particular, assume a player is following a specific deterministic rule, which, at some point in time, prescribes
a change to a different action. Then we postulate that the player will follow the rule with high probability, but with some small probability he might not do so. We define this specification as *noise at the margin* and we shall motivate it in terms of inertia, on the part of players, towards changes in actions.

\[
\Pr(a_t(\omega) = i \mid a_{t-1}(\omega) = i) = 1 \quad \text{if} \quad a_{t-1}(\omega) \in \mathcal{R}(\theta_t)
\]
\[
\Pr(a_t(\omega) = l \mid a_{t-1}(\omega) = i) > 0 \quad \forall l \in A \quad \text{if} \quad a_{t-1}(\omega) \notin \mathcal{R}(\theta_t)
\]

The second rule we analyze formalizes the fact that when players implement their decisions, they sometime make mistakes, that is with some positive, though small probability, they choose a different action. In this specification, we shall assume that mistakes occur in a way that is independent across players, over time and of any payoff consideration. As a result, if it is player \( \omega \)'s updating turn, he will choose:

\[
\Pr(a_t(\omega) \in \mathcal{R}(\theta_{t-1})) > \Pr(\hat{a}_t(\omega) \notin \mathcal{R}(\theta_{t-1})) > 0 \quad \text{for all} \quad \hat{a}_t(\omega) \notin \mathcal{R}(\theta_{t-1})
\]

The rule postulates that each player will *de facto* adopt a mixed strategy that has full support, and that assigns a higher probability coefficient to the action prescribed by the rule. Rules such as the *noisy best reply*; fall into the latter category.

The last behavioral rule we analyze is based on the idea that the probability with which a player adopts each of the available actions is proportional to the expected payoff achievable by choosing that action:

\[
Pr(a_t(\omega) = i \in A) \propto \frac{\exp[\sigma \pi_{\omega}^i(\omega)]}{\sum_{i \in A} \exp[\sigma \pi_{\omega}^i(\omega)]}
\]

The rule is not meant to model any specific behavioural assumption\(^4\), but in-

\(^4\)Though this can be done in terms of conditional logit specification, as in McFadden (1974).
stead formalizes the idea that the ratio between the probability with which a player chooses any two actions is proportional to the difference in the expected payoff achievable. This is clear if we re-write the rule as follows:

\[
\frac{\ln \Pr(a_t(\omega) = i \in A)}{\ln \Pr(a_t(\omega) = j \in A)} \propto \pi^i_\omega - \pi^j_\omega
\]

If we are to motivate probabilistic behaviour in terms of mistakes, the above rule postulates that the probability with which they occur is payoff-dependent, in that very costly mistakes are less likely to occur than relatively less costly mistakes.

Though the stochastic rules we described share some common characteristics, the dynamics they generate might not coincide.

### 2.2.2 Asymptotic behavior

Given any of the above rules, in the noiseless or in the noisy version, the dynamics with which aggregate behaviour evolves is described by \( \theta \), which ranges over the space of all configurations of actions \( \Theta \). We denote the probability that the system be in state \( \theta \) at time \( t \) by \( P_t(\theta) \). Given that all underlying rules specify actions to be taken exclusively in terms of the configuration of play in the previous period, the probability that the system moves from state \( \theta \) at time \( t \) to \( \theta' \) will be time independent and denoted by \( P(\theta' | \theta) \). It follows that:

\[
P_{\theta'}(t) = \sum_{\theta \in \Theta} P(\theta' | \theta) P_\theta(t - 1)
\]

or, in matrix notation:

\[
P(t) = \mathcal{P}P(t - 1)
\]

\(^5\text{The superscript } R \text{ is dropped for notational convenience.}\)
where \( P(.) \) are vectors of \( \Theta \) components and \( \varphi \) is the transition matrix. Given that \( \sum_\theta P_\theta(t) = 1 \) and \( \sum_\theta P(\theta | \theta') = 1 \), (2.1) becomes:

\[
P_\theta'(t) = P_\theta'(t - 1) + \sum_{\theta \neq \theta'} [P(\theta' | \theta)P_\theta(t - 1) - P(\theta | \theta')P_\theta(t - 1)]
\]  \hspace{1cm} (2.3)

(2.3) states that the change in the probability that the system be in state \( \theta' \) between \( t - 1 \) and \( t \) is composed of two parts: an increase due to the transitions from other states into that state and a decrease due to transitions out of this state into other states.

Since (2.2) is linear, it can be iterated to obtain the distribution function \( P \) at time \( t \) as a function of the initial distribution:

\[
P(t) = \varphi^t P(0)
\]  \hspace{1cm} (2.4)

We are interested in the limit of (2.4) as \( t \to \infty \) as a prediction of what will happen to the aggregate system in the long run. Clearly, if \( \varphi \) is known, the above distribution function can always be computed for an arbitrary initial condition. However, in some particular cases the process is such that the limit distribution does not depend on the initial condition, meaning that, for any initial condition, as \( t \to \infty \), \( P_\theta(t) \) converges in probability to a given \( P(\theta) \). In the remaining of this work, a process that has the latter property will be defined to be ergodic.

In the last part of this work, we shall take the view that ergodicity properties of a model make the latter a good predictor of the long run behaviour of the system. In some cases, we shall be able to explicitly derive the invariant distribution and to characterize it as a function of the modeled behavioural rule that generates it. We will then proceed to compare different limit distributions describing the same model, when the parameters that specify the underlying rule take different values.
3.1 Introduction

In this Chapter we focus on the evolution of play under myopic best-reply dynamics. The dynamics rely upon the assumption that players hold static expectations on the environment they are called to interact in, and, whenever they are able to adjust their strategies, they act so as to maximize their expected payoff. The simple formulation therefore combines aspects of fully rational behaviour, with limited information, on the part of players, about opponents’ choices.

In a framework that focuses on the evolution of play over time, myopic best reply dynamics are particularly appealing in that, by construction, whenever the environment is stationary, players’ behaviour is optimal. An immediate implication is that every steady state of the dynamics is an equilibrium. However, cyclic behaviour might occur instead.

In what follows, we will stick to the model introduced in Chapter 1, and, as there, we will consider a general specification of the matching technology, as well
as a general specification of the underlying game. As stated in Section 3.2, the process is a Markov chain. In Section 3.3 we focus on the equivalence classes of the stochastic dynamics. In particular, we identify conditions on the underlying game, sufficient to guarantee convergence to an absorbing state under uniform population matching and under uniform local matching. In Section 3.4 we will then analyze the asymptotic behaviour of the process. If the initial distribution is uniform, for some classes of games we are able to relate absorption probabilities for a uniform population matching model to more basic features of the underlying game.

The model we analyze in this Chapter is analogous to those analysed in Kandori, Mailath and Rob (1993) and Ellison (1993), if mutations are ruled out. More precisely, in these papers three hypothesis describe players' behaviour: (i) inertia, (ii) myopia and (iii) mutations. While we assume that players hold static expectations as to their opponents' choices (myopia hypothesis) and also that, in an asynchronous dynamics, not all players react simultaneously to the environment (inertia hypothesis), we do not postulate any random flow of mutations that hits the system and shapes each player behaviour. The class of games we analyze, as well as the specification of the interaction pattern do, however, account for the aforementioned works as special cases. We see the study of this model as being complementary, rather than alternative to that of the works we quoted. In particular, we believe that considering the questions addressed in this Chapter is essential if the framework of Kandori, Mailath and Rob (1993) or Ellison (1993) is to be generalized.


We recall that in Kandori et al. (1993) the focus is on 2-2 games, and the locally interactive model considered in Ellison (1993) is the particular case where players are located on a circle and they interact only with their nearest neighbours.
As for the specification of the locally interactive setting, the model differs from that analyzed in Berninghaus and Schwalbe (1993), where the focus is on a 2-2 game and the model is entirely deterministic\(^3\). For essentially the same reasons as those stated in the previous paragraph, the model is different from that studied in Blume (1993)\(^4\).

### 3.2 The Model.

The behavioural rule that forms the basis for the learning model we analyze in this Chapter is based on the idea that each player chooses an action that maximizes his expected payoff. His expectation as to what his potential opponents are going to play is shaped by the idea that future play will not be different from what was observed in the recent past. Analogous to the notation previously introduced, let \(\mu_A(\omega)\) be the vector of frequencies of actions observed by player \(\omega\). Each entry is given by:

\[
\mu_j(\omega) = \sum_{s \in \Theta^w: a(s) = j} \mu(s, \omega)
\]

that is the relative frequency with which \(\omega\) observes action \(j \in A\). An action that maximizes \(\omega\)'s expected payoff is then:

\[
a(\omega) \in \text{Arg max}_{a \in A} \sum_j \pi_{ij} \mu_j(\omega)
\]

We assume that, once an action is chosen, a player can only revise his choice when he gets an updating opportunity. As we shall see below, the full specification of the model depends on whether all players receive an updating opportunity simultaneously or if, at each period, only one player does.

\(^3\) As a result of these two assumptions, the analysis is carried out in terms of discrete iterations of Boolean functions.

\(^4\) The model we analyze in Chapter 6 is closer to Blume's model.
The population dynamics is determined by the information that each player has at his disposal when taking a decision. This may, or may not, be the same for all the players. On the basis of this simple distinction, we will study two tractable approximations of the dynamics. The former, synchronous dynamics, relies on the implicit assumption that all the players possess exactly the same information as to the relevant variables. Information as to how the game was played in the past period(-s) of play is hence public domain, and on that basis, all the players simultaneously choose actions. A second specification we adopt, more likely to capture the inherent stochastic nature of information gathering, allows for differences among players in terms of what they know about past play. One way of achieving this, is to postulate an order of updating among players and assume that, within a unit time interval, only one player is allowed to revise his action choice. Thus two players that consecutively update their actions, have information sets that differ by the action of at most one player. We will find it convenient to assume that the sequence of updating is random; in this case in fact the dynamics under asynchronous updating does not depend on the specific order of players' choices.

For the purposes of our work, we will refer to \( \varphi^s \) and \( \varphi^a \) as the transition matrices that describe the dynamics under synchronous and asynchronous myopic best reply, with the convention that we drop the superscript whenever a statement holds for both dynamics. Clearly \( \varphi^s \) and \( \varphi^a \) range over \( \Theta \). Each entry, \( P(\theta' | \theta) \), provides the probability with which the system moves from state \( \theta \) to state \( \theta' \). Given that we know the rule followed by each player when choosing an action (which is time-independent), as well as the order with which decisions are taken in the population, once the subset of past observations players observe is specified, we can fully characterize the process that governs the population dynamics. For simplicity, we assume that, at time \( t \), players only observe action chosen at time \( t - 1 \), and we use a superscript or subscript to explicit the reference to time.

**Definition 7** The transition matrix for the synchronous myopic best-reply dynamics
is denoted as \( p^a \). Each entry is such that for all \( \omega \):

\[
a_{\theta_i}(\omega) \in \operatorname{Arg} \max_{i \in A} \sum_j \pi_{ij} \mu_j^{\theta_i-1}(\omega)
\]

**Definition 8** The transition matrix for the asynchronous myopic best-reply dynamics is denoted as \( p^a \). If states \( \theta \) and \( \theta' \) differ by more than one single player's action, then \( P^a(\theta' | \theta) = 0 \). If states \( \theta \) and \( \theta' \) differ by exactly player \( \omega \)'s action, then \( P^a(\theta' | \theta) > 0 \) and

\[
a_{\theta_i}(\omega) \in \operatorname{Arg} \max_{i \in A} \sum_j \pi_{ij} \mu_j^{\theta_i-1}(\omega)
\]

Finally, \( P^a(\theta' | \theta) > 0 \) if for at least one player \( s 
\)

\[
a_{\theta_i}(s) = \operatorname{Arg} \max_{i \in A} \sum_j \pi_{ij} \mu_j^{\theta_i-1}(s)
\]

The Remark that follows describes the stochastic process generated under myopic best-reply dynamics.

**Remark 5** \( p^s \) and \( p^a \) are finite Markov chains over \( \Theta \).

**Proof.** By the definition of the processes. Note that, given that transition probabilities in \( p^a \) do not depend on the specific sequence of updating, the two processes are Markovian with respect to the same time scale. ■

Two remarks conclude this Section. First, note that transition probabilities depend explicitly on the interaction pattern \( \mu \): if matching is local, a player's best-reply is defined on the basis of the actions adopted in his neighbourhood. As a result, transition probabilities as well will depend only on the relevant subset of players. Second, myopic best-reply dynamics under synchronous or asynchronous updating may behave differently. The intuition is not difficult to describe: assume the system is in state \( \theta \) and for all players \( \omega \), action \( i \) is the unique best reply to \( \theta \). Then, by definition, under synchronous updating the system will move with probability one to
the state \( \theta' \) where everybody in the population is adopting action \( i \). If updating is sequential instead, the process will move through a sequence of different states, that differ from one another by the action of a single player. If action \( i \) is the unique best reply to all these states, then the system will end up in the state \( \theta' \) where everybody in the population is adopting action \( i \). Only in this case, the two dynamics would necessarily show the same path.

### 3.3 The Recurrent Communication Classes.

The aims of this and the next Section are twofold: (a) to study the limit behaviour of the dynamics of a population game \( \Gamma \) under myopic best reply, (b) to relate the results to some more easily detectable features of the underlying game \( G \). The relevance of (b) stems from the fact that in many cases of interest, the underlying game is more tractable an object than the process itself.

We start by studying the equivalence classes of the process. Recall that two states belong to the same equivalence class if they "communicate", i.e. if the process can go from one state to the other. The resultant partial ordering shows the possible directions in which the process can proceed. The minimal elements of the partial ordering of equivalence classes are called ergodic sets, i.e. sets that, once entered, cannot be left by the dynamics. Ergodic sets that contain only one element are called absorbing states. Given a population game, \( \Gamma \), let \( E^p(\Gamma) \) be the set of ergodic sets of the process under \( \rho \) and \( A^p(\Gamma) \subseteq E^p(\Gamma) \) be the set of absorbing states, where \( \rho \) defines transition probabilities under myopic best reply. Given that the state space is finite, \( E^p(\Gamma) \neq \emptyset \).

The next result partially characterizes the set of the absorbing states of the dynamics under myopic best reply, in terms of the equilibria of the population game, studied in Chapter 1.

\footnote{The "A" that follows is not to be confused with the action space, previously denoted as \( A \).}
Remark 6 Given $\Gamma$, let $\Theta^*(\Gamma)$ be the set of strict equilibria of $\Gamma^6$. Then $A_p^*(\Gamma) = A_p^*(\Gamma) = \Theta^*(\Gamma)$.

Proof. A state $\theta$ is absorbing iff. $P(\theta | \theta) = 1$; the mere definition of transition probabilities in $p^a$ and $p^s$ hence proves the assertion.

The equivalence between absorbing states under synchronous and asynchronous myopic best reply does not extend to ergodic sets that contain more than one state, as the next Example show.

Example 9 Consider the following game, usually referred to as anti-coordination game:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

where actions are labelled $T$ and $B$. Assume a population of $\Omega$ players interact in $\Gamma = (\Omega, A, \mu_U)$. Suppose that at time $t$, the system is in state $\theta_T$ where everybody is playing $T$. Clearly, for each player, the unique best-reply is to play $B$. If updating is simultaneous, then with probability one, the system will move to $\theta_B$, where everybody plays $B$. From $\theta_B$ at time $t+1$, for analog reasons, the system will move to $\theta_T$ at time $t+2$, and cycle forever after. Suppose instead updating is sequential. Then from $\theta_B$ at time $t$, the system will move with probability one to a state $\theta'$ where only one player plays $B$, and then with probability $1 - \frac{1}{\Omega} < 1$ to a state where two players are playing $B$, and so on until the fraction of players in the population playing $B$ gets close to $\frac{1}{2}$. At that point two things can happen, according to whether $\Omega$ is even or odd. In the former case, the system will get absorbed in a state where exactly a half of the population plays $B$. In the latter the system will cycle around the two integers around $\frac{1}{2}$. Ergodic sets under the two updating rules would hence be different. Note that this

---

6 An equilibrium is strict if Definition (5) holds with the strict inequality sign for each player.
consideration extends to less trivial games. Consider for example the following game:

\[
B = \begin{pmatrix}
3 & 2 & 2 \\
1 & 0 & 5 \\
1 & 4 & 0
\end{pmatrix}
\]

where actions are labelled T, C and B. Under uniform population matching, the myopic best reply dynamics, admits a unique absorbing state, where all players in the population adopt action T. However, if there are no Ts in the population, then the synchronous dynamics as well as the asynchronous dynamics will cycle around the mixed strategy equilibrium of B, \( p = (0, 5/9, 4/9) \), as a situation entirely analogous to that of the previous example would arise. One would be led to think that a symmetric equilibria in mixed strategies for the underlying game, would work as an attractor for the population dynamics. This, however, is not necessary, as the next example shows.

\[
B' = \begin{pmatrix}
3 & \frac{21}{9} & \frac{21}{9} \\
\frac{21}{9} & 0 & 5 \\
\frac{21}{9} & 4 & 0
\end{pmatrix}
\]

In this case \( p = (0, 5/9, 4/9) \) would fail to be a Nash-equilibrium of \( B' \) (because T becomes the best-reply to \( p \)). However, in this case the synchronous dynamics would still show cyclic behaviour, between the state where everybody is adopting C and the one where everybody is adopting B. The asynchronous would not, in that while moving between the latter two states, a state to which T is the (unique) best-reply would be encountered. Under asynchronous dynamics in fact, the system transits to a state where the frequencies with which actions adopted in the population reproduce a probability distribution over actions that still assigns probability zero to T, but towards which T is optimal.

Given the wide variety of limit behaviour that one could encounter for a generic
underlying game, it is useful to characterize ergodic sets of the population dynamics in terms of basic properties of the underlying game $G$. Trivially, if the game admits a symmetric strict Nash-equilibrium, then a state where all players adopt that equilibrium action is an absorbing state of the dynamics.

**Remark 7** Given $\Gamma = (\Omega, G, \mu)$, $\hat{N}^s(G) \neq \emptyset \Rightarrow A^p(\Gamma) \neq \emptyset$.

The reverse implication does, in general, not hold, in that as we saw in Chapter 1, not all equilibria of the population game reproduce a Nash-equilibrium in terms of the frequencies with which actions are adopted.

As for ergodic sets of a more general nature, the identification relies on the explicit analysis of the dynamics. In what follows we shall analyze models where matching is uniform. Recall, from Definition 1 in Chapter 1, that matching is uniform if $\mu$ is constant over its support. The latter may or may not include all possible ways to couple players in the population. We defined the former case as *uniform population* matching and we referred to it as $\mu_U$. By analogy, a model of *uniform local* matching is defined as follows:

**Definition 10** *The random matching is locally uniform, and referred to as $\mu_L$, if for each $\omega \in \Omega$ and for all $s \in \Omega_\omega$, $\mu(s \mid \omega) = \frac{1}{\# \Omega_\omega}$ and $\# \Omega_\omega = \# \Omega - 1$.*

It is clear that the main difference between a model of *uniform population* matching and a model of *uniform local* matching, is simply the size of $\Omega_\omega$.

We find it useful to analyze the dynamics by looking at a particular statistics of the state $\theta$. We take the latter to be the *average expected payoff* in the population when the configuration of play is $\theta$; we obtain it by aggregating the expected payoff from a round of interaction over all players:

$$\Pi(\theta) = \sum_{\{\omega, s\}} \pi(a_\theta(\omega), a_\theta(s))\mu(\{\omega, s\})$$

$$= \sum_{\omega \in \Omega_\omega} \sum_{s \in \Omega_\omega} \pi(a_\theta(\omega), a_\theta(s))\mu(s \mid \omega)\mu(\omega)$$

53
If matching is uniform, $\mu(\{w, s\})$ is constant, and:

$$\Pi(\theta) \propto \sum_{w \in \Omega} \sum_{s \in \mathbb{F}_w} \pi(a_\theta(w), a_\theta(s))$$

reflecting the fact that each player has exactly the same probability to take part in the interaction. The difference between a model of uniform population matching and a model of uniform local matching is only in the factor of proportionality$^7$. The question we may ask then is what happens to this quantity as the myopic best-reply dynamics unfolds. It is clear that if the system moves from $\theta$ to $\theta'$ under asynchronous updating, the expected payoff for the player who switches to his best reply in $\theta'$ is higher than in $\theta$. His move will affect only the expected payoff of the players with whom he interacts (all other players in a population matching model, or only his neighbours in a locally interactive setting). The sign of the latter effect is however unclear. The lemma that follows identifies conditions on the payoffs of the underlying game, sufficient to guarantee that at each step of the asynchronous myopic best reply dynamics, the average payoff in the population will be strictly increasing. We let $\Gamma = (\Omega, G, \mu_U) = \Gamma_U$ and $\Gamma = (\Omega, G, \mu_L) = \Gamma_L$. As previously done, we denote by $\mu_\theta^i(\omega)$ the frequency with which player $\omega$ observes action $i$ within the set of his potential opponents. If $\#\Omega_\omega^i$ is the number of $\omega$'s potential opponents choosing action $i$, then $\mu_\theta^i(\omega) = \frac{\#(\omega)}{\#\Omega_\omega}$ under uniform local matching and $\mu_\theta^i(\omega) = \frac{\#(\omega)}{\#(\omega)+1}$ under uniform population matching.

**Lemma 11** In asynchronous myopic best-reply dynamics for a population game $\Gamma_U$ or equivalently $\Gamma_L$, suppose player $\omega$ switches from action $j$ in state $\theta$ to action $i$ in state $\theta'$ and suppose the latter is his strict best-reply. Then,

$$\sum_{i \in A}(\pi_{ij} - \pi_{ij})\mu_\theta^i(\omega) > 0 \Rightarrow \Delta \Pi = \Pi(\theta') - \Pi(\theta) > 0$$

$^7$Under $\mu_L$, as previously defined, $\mu(s \mid \omega) = \frac{1}{\#(\omega)}$, $\mu(\{\omega, s\}) = \frac{2}{\#(\omega)+1}$ and $\mu(\{\omega\}) = \sum_{s \in \mathbb{F}_\omega} \mu(\{\omega, s\}) = \frac{2}{\#(\omega)+1}$ and $\mu(\{\omega\}) = \frac{2}{\#(\omega)+1}$ respectively.

54
Proof. Simple algebra shows that:

\[
\Delta \Pi \propto \sum_{\omega \in \Omega} \sum_{s \in \Omega_s} [\pi(a_{\theta'}(\omega), a_{\theta}(s)) - \pi(a_{\theta}(\omega), a_{\theta}(s))] = \\
= \sum_{s \in \Omega_s} [\pi(a_{\theta'}(\omega), a_{\theta}(s)) - \pi(a_{\theta}(\omega), a_{\theta}(s))] + \\
+ \sum_{s \in \Omega_s} [\pi(a_{\theta}(s), a_{\theta'}(\omega)) - \pi(a_{\theta}(s), a_{\theta}(\omega))]
\]

where the first relation and the first equality stem from the definition itself; the
second equality expresses the first in terms of the payoffs of the underlying game. The
strict inequality stems from the fact that, by assumption, \( a_{\theta'}(\omega) < a_{\theta}(\omega) \).
Given that the constant of proportionality is positive by assumption, this is enough
to prove the assertion. 

We now remark on the above result, that, as can easily be noticed, relies crucially
on the assumption that matching is uniform and not on the distinction between local
and population matching. First, the condition is clearly not necessary. Informally,
it requires player \( \omega \) to find it optimal to choose action \( i \) over action \( j \), given the
configuration of play among his potential opponents, also if he had to play a different
game, where the payoff matrix is exactly the transpose of the original one. For later
reference, we define a game that has this property to be balanced:

**Definition 12** A symmetric normal form game, \( G \), is defined to be balanced if its
payoff matrix is symmetric, i.e. if \( \pi_{ij} = \pi_{ji} \) for all \( j \neq i \).

Second, the result relies on the fact that player \( \omega \) switches to his strict best-reply,
and some games do not allow this possibility, because for example their set of strict
Nash-equilibria is empty. Moreover, the underlying game may admit Nash-equilibria
in mixed strategies, in which case a player may have multiple best-replies among
which to choose. For reasons that will become clear in what follows, we would like
to focus on population games for which this possibility does not arise. We do that by exploiting the fact that the model of local uniform matching we analyze is very general, and allows for an arbitrary choice of the size of the neighbourhood. In words, we require the size of the neighbourhood to be such that observed frequencies of play do not reproduce any of the mixed strategy equilibria of the underlying game. Formally we state the following Assumption:

**Assumption 4** \( \Gamma_L \) is such that for all players \( \omega \) and all configurations of play \( \theta \), there exists an \( i \in A \) for which:

\[
\text{Argmax}_{j \in A} \sum_{i \in A} \pi_{ji} \mu_i^\theta(\omega) = \{i\}
\]

Assumption 4 is to be taken as a compound assumption on the population game. If the underlying game admits a mixed strategy Nash-equilibria with support given by the set \( m^* \), then the assumption requires to choose an \( n = \|\Omega_\omega \| \) such that, for all \( i \) and \( j \neq i \) in \( m^* \)

\[
\sum_{i \in A} (\pi_{ii} - \pi_{ji}) \mu_i(\omega) \neq 0
\]

where, we recall, \( \mu_i(\omega) = \frac{\#\Omega_i}{n} \). We shall refer to a population game that satisfies the above Assumption as a population game with no ties\(^8\).

Lastly, we note that the average expected payoff function is bounded. As it is intuitively clear, an absorbing state where all players adopt the strategy corresponding to a strict Nash-equilibrium of the underlying game constitutes a local maximum of the average expected payoff function. In general, the converse is not necessarily true: though the function is maximized, players might have multiple best-replies, and adopt either of them with positive probability. Again, if the game has no ties, this caveat does not apply.

These considerations explain the next result.

\(^8\)Clearly, a particular case where ties cannot be ruled out in this fashion is if the underlying game is such that \( \pi_{ii} = \pi_{ji} \) for all \( i, j \) and \( l \). We shall disregard this case.
Theorem 13 If $\Gamma$ admits no ties (i.e. if it satisfies Assumption 4) and $G$ is balanced, then $A^\pi(\Gamma) = E^\pi(\Gamma)$.

Proof. It is obvious to notice that if $G$ is balanced, $\pi_i = \pi_u$ for all $(i, u)$. Hence the condition of which in Lemma 11 is automatically satisfied. Thus, whenever a player switches to his strict best-reply, the average expected payoff in the population is strictly increasing. Given that ties are ruled out by assumption (and by an appropriate choice of $n = \#W$), given an initial condition at time $t_0$, along any asynchronous path of a myopic best-reply dynamic, the average expected payoff is monotonically increasing. Since the function is bounded, it will reach a local maximum at some finite time $t$. By definition, the value of the function cannot be increased any further in a single step. This can only be so if all players are adopting their strict best-reply to $\theta_i$ (otherwise, with strictly positive probability a player would receive an updating opportunity and change action). But this coincides with the definition of an absorbing state of the process, from which the assertion follows.

As previously noted, the condition stated in Lemma 11, as well as that of Theorem 13 are not necessary in order for convergence to obtain. De facto, if a game is balanced, the standard notion of average payoff of a game $G$ as $\sum_{i=1}^n \sum_{j=1}^n p_i p_j \pi_{ij}$ can be taken to represent a potential or an energy function of the system: at each step of the dynamics, with positive probability, the system "climbs up" towards a local maximum. If the population game has no ties, then all these transitions take place with probability one, and the system reaches an absorbing state in correspondence of a local maximum.

Though the result of Theorem 13 is very general, not all games are balanced. For example, it is not necessarily satisfied in a Coordination game. In a 2-by-2 coordination game for which $0 < \pi_{11} - \pi_{22} < \pi_{21} - \pi_{12}$, such as a Stag-hunt game, it is easily seen that it fails to hold. However, convergence is obtained, as the next result shows.
Theorem 14  

(a) Given \( \Gamma = (\Omega, C_m, \mu_U) \), if \( \pi_{ii} - \pi_{ii} > \pi_{ij} - \pi_{ij} \quad \forall j \neq i \neq l \), then \( A^p(\Gamma) = E^p(\Gamma) \).

(b) Given \( \Gamma = (\Omega, C_2, \mu_L) \), if \( \Gamma \) has no ties (i.e. if it satisfies Assumption 4), then \( A^p(\Gamma) = E^p(\Gamma) \).

Proof. Under uniform population matching (i.e. in case (a)) the results is obvious: following Remark 4 in Chapter 1, given a state \( \theta \) the best-reply to \( \theta \) is uniquely identified and the same for all players\(^9\). If updating is asynchronous, the system will move with probability one, from any state \( \theta \) to a state \( \theta' \) where one more player adopts the action that is the (unique) best-reply to \( \theta \). Hence again the system will converge to a state where all players in the population adopt the same action, and the latter is necessarily absorbing. We now focus on case (b), i.e asynchronous updating and local matching.

It is clear why the above logic fails to hold in a locally interactive system: if action \( i \) is a best-reply for player \( \omega \), who basis his choices on the configuration of play in \( \Omega_\omega \), there is no a priori reason to believe that the same might hold for player \( s \), who instead sees \( \Omega_s \). \( \Omega_\omega \) and \( \Omega_s \) might not have anything in common, or only be partially overlapping. But if \( \Gamma \) has no ties, because \( \|\Omega_\omega \) is chosen such that the (unique) mixed strategy equilibrium of \( G \) cannot be reproduced by observed frequencies of play, then we are about to show that, with strictly positive probability the system reaches an absorbing state. Suppose the system starts in a non-absorbing state \( \theta \), where there are \( i \)-players and \( j \)-players; some of each group (not all) are best-responding, some other are not. With positive probability a player who is adopting \( i \), but is not best-responding, will receive an updating opportunity and will switch to action \( j \). The system then moves to \( \theta' \). Again, with positive probability a player who is adopting \( i \), but is not best-responding, will receive an updating opportunity and will switch

\(^9\)Recall that, by assumption, \( \|\Omega > 2 \). As a curiosity, the synchronous dynamics of a 2-players coordination game might show cyclic behaviour, between the two states where players choose different actions. In Chapter 4 we will use an example analog in content, to show that synchronous updating may lead to cyclic behaviour in a locally interactive setting.
to action \( j \). The system then moves to \( \theta'' \). Given that the population is finite, in this fashion, the system reaches, with strictly positive probability a \( \theta \), where either there are no \( i \)-players left, or all \( i \)-players in the population are best-responding. Suppose the latter case occurs (in the former, the proof is ended). Then with positive probability a \( j \)-player who is not best-responding will be chosen to update his action choice, and will conveniently choose action \( i \). By this doing he will not "harm" any potential \( i \)-player, in that the relative advantage of action \( i \) over action \( j \) is (positive, from \( \theta \) and) strictly increasing in the observed frequency of \( is \). One at a time all non best-responding \( j \)-players will, with positive probability, switch to \( i \). Hence, the system will reach a state where either there are no \( j \)-players left or all \( i \)-players and all \( j \)-players are adopting their strict best-reply. In both cases the system has reached an absorbing state. Given that \( \theta \) was chosen arbitrarily, the result follows\(^{10}\).

It is intuitively clear why the logic followed in the last part of the above proof does not extend to a generic \( m-m \) Coordination game. The system, starting from a state where all \( i \)-players are best responding, moves with strictly positive probability to another state where a non best-responding \( j \)-player switches to action \( l \neq i \). But while under the stated assumption on the underlying payoffs, the relative advantage of action \( i \) is strictly increasing in the observed frequencies of \( i \)-players\(^{11}\), it is also strictly decreasing in at least one other action, say \( l^2 \). Hence whenever a \( j \)-player switches to action \( l \), this might undermine the optimality of the choice made by an \( i \)-player, and by this lead to a vicious circle in the logic.

In the Section that follows we will explicitly study the time evolution of the stochastic process under myopic best reply.

---

\(^{10}\)Clearly, in this general framework, absorbing states cannot be explicitly characterized. We shall get back to this point in Chapter 4.

\(^{11}\)To see this it suffices to notice that if \( p \) is any probability distribution over \( A \), \( \frac{\partial}{\partial p_i} \sum_j (\pi_{ij} - \pi_{ij})p_j = (\pi_{ii} - \pi_{ii}) - (\pi_{ij} - \pi_{ij}) > 0 \).

\(^{12}\)in that \( \frac{\partial}{\partial p_i} \sum_j (\pi_{ij} - \pi_{ij})p_j = (\pi_{ii} - \pi_{ii}) - (\pi_{ii} - \pi_{ii}) < 0 \).
3.4 Asymptotic Behaviour and Likelihood of Equilibria.

The results that follow do not depend on any explicit characterization of the ergodic sets of the process. In order to avoid heavy notation, we will therefore refer to a generic ergodic set of the process as a single ergodic state. Formally, if the process, \( \mathcal{P} \), admits ergodic sets different from absorbing states, we will implicitly look at a different transition matrix, say \( \mathcal{P}' \), where all states belonging to the same ergodic set have been lumped into a single state. This is legitimated by the fact that, by definition, ergodic sets cannot be left by the dynamics. As a result, \( \mathcal{P}' \) is an absorbing Markov chain, where the behaviour of the process before absorption is exactly the same as the original process before hitting an ergodic set for the first time.

Furthermore, we find it useful to refer to a version of \( \mathcal{P}' \), where states have been rearranged, such that the first \( \#\Theta - s \) states are absorbing and the remaining \( s \) states are transient:

\[
\mathcal{P}' = \begin{pmatrix} I & O \\ R & Q \end{pmatrix}
\]

\( I \) is a \( (\#\Theta - s) \times (\#\Theta - s) \) identity matrix that refers to the absorbing states of the new process; \( O \) is an \( s \times (\#\Theta - s) \) null matrix; \( R \) and \( Q \) are \( (\#\Theta - s) \times s \) and \( s \times s \) respectively, and concern the behaviour of the process before an absorbing state is reached. For notational convenience, we refer to this new process as \( \mathcal{P} \) in what follows, to the set of absorbing states as \( S \) and to its complement in \( \Theta \) as \( \bar{S} \).

Given a probability distribution over initial conditions, a row vector \( P(0) \), the evolution of the system can be described (as in (2.3) in Chapter 2) by the recursive equation:

\[
P(0)\mathcal{P}' = P(t)
\]  

(3.1)

Investigating the limit behaviour of the process amounts to studying the asymptotic behaviour.
totic behaviour of (3.1), that is to characterize the \( \lim_{t \to \infty} P(t) \). In general, this limit will depend on the behaviour of the powers of \( \varphi \), as well as on the initial condition.

To see this, recall that from Remark 6 we know that there is a one-to-one correspondence between the set of absorbing states of the dynamics and the set of strict equilibria of the population game. In Chapter 1 we analyzed the relation between these equilibria and the set of equilibria of the underlying game, and we found that for a given population game each equilibrium corresponds to a correlated equilibrium of the game being played. Given that the stochastic process defining the myopic best-reply dynamics is a finite Markov chain, we are led to infer that \( \varphi \) has an entry equal to one for each of the latter. A trivial exception is when the process admits a unique ergodic set. By the standard result we report below, the process will then be ergodic and independently of the initial condition will converge to that set.

**Theorem 15**  Given \( \varphi \), if \( A^\varphi(\Gamma) = \{\emptyset\} \), then for any \( P(0) \)

\[
\lim_{t \to \infty} P(t) = \begin{cases} 
1 & \text{for } \emptyset \\
0 & \text{for all } \emptyset' \neq \emptyset
\end{cases}
\]

**Proof.** See, for example, Kemeny and Snell (76), Theorem 6.2.1.

Whether this is the case, can often be established by looking at the underlying game, as stated next.

**Theorem 16**  Given \( \Gamma = (\Omega, G, \mu) \), if there exists an \( i \in A \) such that \( \pi_{ij} > \pi_{ij} \quad \forall l \neq i \), then \( A^\varphi(\Gamma) = \{\emptyset_i\} \).

**Proof.** If the game admits \( i \) as a strictly dominant strategy, then the unique correlated equilibrium of \( G \) assigns probability one to \((i, i)\). The corresponding unique absorbing state of the process is then when all players play action \( i \).

As Example 9 shows, the condition is only sufficient.
In case of multiplicity of absorbing states, the process admits multiple limit distribution\(^{13}\). Given that \(\varphi\) provides us with detailed information about absorption probabilities for each absorbing state, we might want to "rank" absorbing states in terms of the latter. To this aim, let \(P^n(\theta | \theta')\) be the probability that the system, starting from state \(\theta'\), reaches absorbing state \(\theta\) in \(n \geq 1\) steps. Then the basin of attraction of a state, \(\theta \in \Theta\) at time \(t\) is defined as \(B(\theta) = \{ \theta' \in \Theta : P^n(\theta | \theta') > 0 \}\). A state \(\theta'\) hence belongs to the basin of attraction of absorbing state \(\theta\) if it communicates with it. If \(\pi\) is any probability distribution over \(\Theta\), we define \(B_\pi(\theta) = E_\pi[P^n(\theta | \theta')]\) to be the \(\pi\)-size of \(B(\theta)\). The next Remark derives the values of \(P^n(\theta | \theta')\) for the different absorbing states of \(\varphi\) in terms of its entries.

**Lemma 17** Let \(P^*\) be a \(\|\Theta \times \|\Theta\) matrix, where a generic entry is \(P^n(\theta | \theta')\) for \(\theta\) absorbing and 0 otherwise. Then

\[
P^* = \begin{pmatrix}
I & 0 \\
(I - Q)^{-1}R & O
\end{pmatrix}
\]

**Proof.** Kemeny and Snell (1976), Theorems 3.3.7.

Each column vector of \(P^*\) gives the probability of absorption in any of the \(\|\Theta - s\) absorbing states, in terms of the original transition probabilities. For the column regarding absorbing state \(\theta\), the entry corresponding to state \(\theta'\) quantifies \(P^n(\theta | \theta')\) as a function of \(\varphi\). If \(\theta'\) communicates only with absorbing state \(\theta\), then the corresponding entry will be one. In general, it is worth noticing that \(P^*\) itself is a transition matrix, reflecting the fact that \(\bigcup_{\theta \in \Theta} B(\theta) = \Theta\).

Given any \(\pi\), we can then calculate the \(\pi\)-size of the basin of attraction of each absorbing state. This would give us a criterion to "compare" different absorbing states of the process, conditional on \(\pi\), that can be taken to be the initial probability

\(^{13}\)In the terminology that we will introduce in Chapter 6, the process is non-ergodic. Note that, given that the set of limit distributions is convex, whenever there are multiple absorbing states, the number of invariant distributions is infinite.
vector of the process. As it is clear, unless the model is very simple, the explicit
derivation of the transition matrix that governs the dynamics is at best very tedious.
In what follows we will hence relate the analysis of the transition matrix to basic
properties of the underlying game.

In particular, for some classes of games and under specific assumptions on the
initial condition of the process, we are able to identify the basins of attraction of each
absorbing state in terms of the best-reply regions of the underlying game. We define
the best-reply regions as follows:

**Definition 18** Given $G$, the best-reply region of action $i$, $BR(i)$, is the set of probability distributions $q: A \rightarrow [0,1]$ such that $\sum_{j=1}^{m} q_j (\pi_{ij} - \pi_{ij}) \geq 0$ for all $l \neq i$.

Best-reply regions are of interest because they can be represented as subsets of
the $m - 1$ simplex, by solving the system of linear inequalities$^{14}$. Clearly if $(i,i)$ is
a symmetric strict Nash-equilibrium of $G$, its best-reply region will not be empty,
and will contain $q_i = 1$. If the game admits a symmetric Nash-equilibrium in mixed
strategies $p$, we may conventionally extend the definition by allowing $BR(p)$ to include
all the best-reply regions of each action $i$ for which $p_i > 0$.

Each configuration of play is a mapping that assigns an action to each player;
by looking at the relative frequencies with which actions are played, we can always
identify each state $\theta$ with a point in the $m - 1$ simplex, $\Delta_{m-1}$. The state-space of
the process can then be characterized in terms of a finite array of points in $\Delta_{m-1}$.
If transition probabilities depended only on the fractions of players adopting each
action, we could analyze the dynamics visually by looking at the finite approximation
of $\Delta_{m-1}$. However, this is in general not possible: if the population is finite, transition
probabilities depend on each player "view of the world", and not on a single statistics
of the state $\theta$. In what follows, we identify some particular cases where the payoffs

---

$^{14}$Incidentally, we recall that in the economy of this Chapter each player is assumed to be able to
perform such a calculation.
of the underlying game permits such an approximation. Let \( m_L \) be the Lebesgue measure defined over \( \Delta_{m-1} \) and \( \theta_p \) be the configuration of play where all players in the population adopt the same and any of the actions in the support of \( p \).

**Theorem 19** Given \( \Gamma = (\Omega, C_m, \mu_U) \), suppose \( \pi_{ii} - \pi_{ij} > \pi_{ij} - \pi_{lj} \) for all actions \( i \neq j \neq l \). Let \( \hat{\pi} \) be a probability distribution uniform over \( \Theta \). Then \( B_\hat{\pi}(\theta_i) = m_L(BR(i)) \).

**Proof.** Theorem 14 rules out ergodic sets different from absorbing states. These are all, and only those, states where all players adopt the same action (Remark 4 in Chapter 1). Lemma 17 identifies absorption probabilities in terms of the entries of the transition matrix. Specifically, in the matrix \( P^* \) there will be an entry equal to one if state \( \theta' \) communicates with absorbing state \( \theta_i \). This will be the case if in \( \theta' \) there is at least one player adopting action \( i \) as his strict best-reply. Hence, given \( \hat{\pi} \), the relative size of the basin of attraction of \( \theta_i \) is simply the proportion of states in \( \Theta \), towards which \( i \) is the best-reply. Thus:

\[
B_\hat{\pi}(\theta_i) = E_\hat{\pi}(P^n(\theta | \theta')) = \frac{\#B(\theta_i)}{\#\Theta} = \frac{\sum_{k=1}^{\#\Theta} I_{BA(\theta_i)}}{\#\Theta} = m_L(BR(i))
\]

where \( I_{BA(\theta_i)} \) is an indicator function equal to one iff. \( \theta \in BA(\theta_i) \), and states in \( \Theta \) have been indexed by \( k \). □

Though the asymptotic behaviour depends on the initial condition, under uniform matching the above result allows to infer the likelihood with the system is absorbed in any of the different equilibria by looking at the best reply regions of each. Given that the latter are compact sets, this amount to calculating their integral in the \( m - 1 \) simplex. In general whenever the overlap between basins of attraction of different equilibria does not have zero measure, this does not work. Of course the caveat does not apply if the underlying game is \( 2 \times 2 \). If one equilibrium of the underlying game has the greatest best-reply region, we may then consider the configuration of play.
where everybody adopts that equilibrium action as having the maximum basin of attraction when every state is equally likely to be the initial condition of the process.

If the underlying game is a coordination game, the characterization of the relative sizes of the best-reply regions can also be done by ranking different equilibria in terms of the risk associated to each equilibrium action. Recall that, from Harsanyi and Selten (1988), for a $2 \times 2$ coordination game an equilibrium, say $(T, T)$, is defined to be risk dominant with respect to the other equilibrium, $(BB)$, if it is the best reply to a uniform probability distribution over actions, or, equivalently, if $\pi_{TT} - \pi_{BT} \geq \pi_{BB} - \pi_{TB}$. As it is obvious, by construction, the best reply region of such an equilibrium is greater than that of the other equilibrium. Things are not as straightforward for a generic coordination game, for which the above criterion only allows for pairwise comparisons, between different equilibria. The notion of risk-dominance has been extended to 3-by-3 coordination games by Kandori and Rob (1993) (as simultaneous risk-dominance of one equilibrium with respect to the others) and by Ellison (1995) (in terms of $\frac{1}{2}$-dominance, a particular case of the notion of $p$-dominance introduced in Morris, Rob and Shin (1995), according to which a symmetric equilibrium $(T, T)$ is $p$-dominant if it is the strict best-reply to any mixed strategy placing probability at least $p$ on $T$).

We generalize the original definition as follows:

**Definition 20** Given $C_m$ with action space $A$, consider any two actions $i$ and $l \neq i$. Let $p^{i \leftarrow i} : A \rightarrow [0, 1]$ be any probability distribution over $A$, such that $p_i = p_l$. Equilibrium $(i, i)$ is risk-dominant if action $i$ performs strictly better than any action $l \neq i$ with respect to any given $p^{i \leftarrow l}$, i.e. if and only if $\sum_{j=1}^{m} p^{i \leftarrow l}_j (\pi_{ij} - \pi_{il}) > 0$ for all $l \neq i$.

For an underlying 2-by-2 game Definition 20 coincides with the one of Harsanyi and Selten; for a 3-by-3 coordination game the requirement is sufficient to guarantee that those quoted above are fulfilled (but is not necessary); we are not aware of
any generalization for \( m > 3 \). It is not difficult to show that an equilibrium that is risk-dominant has by construction the largest best-reply region:

The above figure shows the best-reply regions of a 3-by-3 coordination game. If equilibrium \( A \) is risk-dominant, as in the above definition, its best-reply region includes the shaded area. The latter is delimited by the line \( bB \) (the set of probability distributions that assign equal probability to actions \( A \) and \( C \)) and \( cC \) (the set of probability distributions that assign equal probability to actions \( A \) and \( B \)).
Chapter 4

Majority Rules
and Noise at the Margin

He who does not expect will not find out the unexpected,
for it is trackless and unexplored.
(VII)

4.1 Introduction

In this Chapter we focus on the population dynamics generated by a particular class of boundedly rational behavioural rules, that we characterize as majority rules. The underlying idea is that when a player is called to choose an action, in a setting entirely analogous to that of Chapter 3, he will simply imitate the most popular action among those that he observes. Clearly, choices made in this way may or may not be optimal in terms of payoff. We see majority rules as constituting a plausible approximation of economic behaviour in settings where there exists an incentive to conform to other players' actions. Coordination games provide a natural formalization of this underlying structure and hence constitute the focus of this Chapter. Besides, we assume players take choices with inertia and provide a rationale in terms of noise
at the margin. The idea is that whenever the learning rule tells a player to change action, then with some positive probability, this will not happen because of a mistake, owing to which the player may choose any other available action. On the contrary, whenever the rule prescribes no changes, the player will follow the prescription. Lastly, we focus on locally interactive systems, where each player interacts only with a subset of players in the population, conventionally referred to as neighbours. Some classes of neighbourhood arrangements will be explicitly characterized.

In the same line of Chapter 3, we will analyze the population dynamics generated by the model. Theorem 14 in Chapter 3 provided convergence results, for coordination games, under myopic best-reply dynamics. The findings relied on the assumption that the interaction structure was such that, for each player, the best-reply could never be a correspondence. One way of interpreting the requirement is to assume that players are only allowed to choose pure strategies and to postulate an arbitrary tie-breaking rule to deal with cases of indifference. We shall keep this assumption in the first part of this Chapter. However, in the second part (namely in Section 4.5), given that we specialize the model further by considering the particular spatial arrangement of a Torus, we are able to relax the requirement.

The Chapter is organized as follows. In Section 4.2 we describe in detail the features of the model and relate the framework to that of Chapter 1. In Section 4.3 we analyze the relation between payoff maximizing behaviour and majority rules. Then, in Section 4.4, we analyze the issue of convergence. Finally, in Section 4.5 we specialize the model further by considering the particular spatial arrangements of a Torus, but a mildly more general majority rule than in previous Sections. We prove convergence for this system and fully characterize absorbing states of the process.

The results of Section 4.5 are entirely analogous to those obtained in Anderlini and Ianni (1996a). Sections 4.3 and 4.4 constitute a generalization of the model of which in Anderlini and Ianni (1996a) to a generic $m$-by-$m$
underlying game. It should be stressed that the way we model imitative behaviour is substantially different from the specifications adopted for example in Schlag (1994) or Binmore and Samuelson (1995), in that in our model imitation takes place irrespective of any payoff considerations. The model is conceptually closer to the literature on "hearing" models or (purely) social learning, revised for example in Gale (1995), although we deal with an explicitly interactive contest, where informational externalities do affect players' payoffs. Besides, it should be stressed that the stochastic component modelled as "noise at the margin" is very different from the "noise" used to model mutations in Kandori, Mailath and Rob (1993) or Young (1993) and many others: while the latter postulates that players adopt de facto a completely mixed strategy, the former requires this to be the case only when the learning rule prescribes a change in action.

4.2 The Model

The model we analyze formalizes a process of learning analog to that of Chapter 3. As we recall, the model was fully specified by a Population game, a behavioural rule mapping information about previous play into the action space, and aggregate population dynamics, defined by the order with which choices were taken in the population. We spell out below the further details of the learning process we analyze in this Chapter.

4.2.1 Local Interaction

We consider a class of systems in which the learning rules that players use are restricted to use information about neighbouring players only. Players can learn from their neighbours only. We interpret the local nature of the learning interaction as a
stronger version of the assumption of limited rationality which underlies the literature on learning, possibly coupled with information gathering and/or computational costs. Perhaps it is possible in principle for the players to take into account what goes on in 'distant' parts of the system, but it is too costly for them to discover this and/or is too complex a task for them do decide how to take the information into account.

Focusing on locally interactive learning has several important analytical consequences. As we remark in Section 4.4, when the learning rule considered is local, convergence may not be obtained even in the simplest cases. As we anticipated in Chapter 1, local learning rules may yield steady states of the system which are radically different from the steady states of the uniform population analog. In Section 4.5 we characterize the steady states of a particular local learning system when the underlying game is a 2-by-2 coordination game. While the uniform population matching model would converge to an equilibrium in which all players take the same action, in the local interaction case we find that both actions available can survive in the long-run. Only local coordination is obtained.

We now move on to the technical aspects. We analyze a population game where matching is locally uniform, as under Definition 10 of Chapter 3. Given $\Gamma = (\Omega, G, \mu_L)$, at times we will find it convenient to think of the elements of $\Omega$ as being the vertices of a graph, where an edge connects two elements if and only if $\mu$ is strictly positive. We shall refer to the latter graph as $\Omega$ itself; for $n = \sharp \Omega$, the assumption of local uniformity then implies that $\Omega$ is $n$-regular$^1$. Since we consider a finite number of players each with a given fixed number of neighbours, we are implicitly assuming away any special 'boundary conditions'. Some examples of spatial arrangements which fit the structure we have just described are the following. A finite set of points on a 'circle'; each point with a left and a right neighbour. The vertices of a cube; each with the neighbours corresponding to the three adjacent edges. A square with a grid

$^1$An $n$-regular graph is an (undirected) graph, in which each vertex is connected to exactly $n$ edges (Andrásfai (1977)).
of horizontally and vertically aligned points which is then folded to form a Torus, so that the east boundary of the square is joined with the west boundary, while the south boundary is joined with the north boundary of the square. We study this special case of the neighbourhood structure at length in Sections 4.5.

4.2.2 Majority Rules and Noise at the Margin.

In this Section we describe the general majority rules which form the basis for the learning rule we shall use throughout this Chapter. Consistently with the notation previously introduced, let \( a_\theta(\omega) \in A \) be the action of player \( \omega \) in state \( \theta \in A^\Omega \). Moreover, let \( \mu_A(\omega) \) be the vector of frequencies of actions that \( \omega \) observes. Specifically each entry \( \mu_j(\omega) \) is the relative frequency with which action \( j \in A \) is observed in \( \omega \)'s neighbourhood. Clearly, given that matching is uniform, the latter is simply given by the number of \( j \)-players out of the \( n \) potential opponents for player \( \omega \). Player \( \omega \) decides what to play on the basis of \( \mu_{\omega}(\omega) \). Each player only observes action chosen in the previous period. We assume that, at time \( t \), each player imitates the action adopted, at time \( t - 1 \), by the relative majority of his neighbors. Given that each player follows exactly the same rule, the aggregate synchronous dynamics are defined as follows.

**Definition 21** The transition matrix for a Majority Rule is defined as \( \Psi^m \). Each entry is such that for all \( \omega \):

\[
a_{\theta i}(\omega) = i \in A \text{ if } \mu_{\theta i}^{\omega t-1}(\omega) \in \max_{j \in A} \mu_j^{\omega t-1}(\omega).
\]

We further postulate that choices are made with inertia. We motivate inertia in updating choices in the following way. Let a particular learning rule be given. Suppose that the prescription which this learning rule yields for a particular player, say \( \omega \), at time \( t + 1 \) is the same as the action that he actually played at time \( t \). We then assume that \( \omega \) will follow the prescription given by the learning rule at \( t + 1 \)
with probability one. Suppose, by contrast, that the given learning rule prescribes that \( \omega \)'s action at \( t + 1 \) should be different from the action he took at \( t \). In this case we assume that \( \omega \) will make a mistake with strictly positive probability. We call this type mechanism for mistakes noise at the margin.

There are two main interpretations of noise at the margin which we introduce below. The first is that experimentation is triggered by change. If a player sees no reason for change then he also sees no reason for experimenting with a new course of action in the underlying game. Whether this is an appealing interpretation of our model below is obviously dependent on the particular economic (or other) interpretation which is given to the system, and ultimately a matter of taste. Whenever the motto 'why fix it if it ain't broken' seems appropriate to the circumstances this interpretation seems correspondingly fitting. The second interpretation is that of inertia. The noise at the margin can be simply interpreted as stipulating that whenever a player's learning rule prescribes a change of action, then with some positive probability inertia will prevail, and he will stick to the action he took in the previous period.

Formally, the aggregate dynamics of a Majority Rule with Noise at the Margin is defined as follows

**Definition 22** The transition matrix for a Majority rule with Noise at the Margin is defined as \( \varphi^n \). Each entry is such that for all \( \omega \):

\[
P(a_{\theta_t}(\omega) = i | a_{\theta_{t-1}}(\omega) = i) = 1 \quad \text{if } \mu_i^{\theta_{t-1}}(\omega) \in \max_{j \in A} \mu_j^{\theta_{t-1}}(\omega)
\]

\[
P(a_{\theta_t}(\omega) = l | a_{\theta_{t-1}}(\omega) = i) > 0 \quad \forall l \in A \quad \text{if } \mu_i^{\theta_{t-1}}(\omega) \notin \max_{j \in A} \mu_j^{\theta_{t-1}}(\omega)
\]

### 4.3 Optimality of Majority Rules

It is clear that majority rules do not explicitly depend on the payoffs of the game being played; as such they are not necessarily consistent with payoff maximization. However, it is also clear that whenever the relative advantage, in terms of expected payoffs, of playing action \( i \) over action \( j \) is increasing in the probability with which
the opponent chooses action $i$, then by conforming to an existing majority, one would also achieve a higher payoff. The Theorem that follows shows that possible underlying games for which majority rules are consistent with payoff-maximizing behaviour are to be found in the class of Coordination games. The Theorem will make use of the following Definition, analogous to Definition 20 in Chapter 3:

**Definition 23** Given $C_m$ with action space $A$, consider any two actions $i$ and $l \neq i$. Let $p^{i=l}: A \rightarrow [0,1]$ be any probability distribution over $A$, such that $p_i = p_l$. Action $i$ and action $l$ are risk-equivalent if they perform equally well with respect to any given $p^{i=l}$, i.e. if

$$\sum_{j=1}^{m} p_j^{i=l}(\pi_{ij} - \pi_{lj}) = 0 \quad \forall p^{i=l}.$$

We note that, consistently with previous remarks, the Definition is entirely equivalent to Harsanyi and Selten's for a 2-by-2 coordination game, and only sufficient for the generalization to the a 3-by-3 case provided by Kandori and Rob (1993) and Ellison (1995).

**Theorem 24** Given is $C_m$ with action space $A$. Majority rules as in Definition 21 are payoff maximizing for a Coordination game $C_m$, if and only if all actions $i$ and $l \neq i$ are risk-equivalent.

**Proof.** Let $B : [0,1] \rightarrow A$ and $M : [0,1] \rightarrow A$ be the two mappings describing the best-reply and the majority rule correspondence respectively. We need to show that if, and only if, $\sum_j p_j^{i=l}(\pi_{ij} - \pi_{lj}) = 0 \quad \forall p^{i=l}$ for any pair of actions $i$ and $l$, then $B^{-1}(A) = M^{-1}(A)$, i.e.:

$$p_k \geq p_h \iff \sum_{j=1}^{m} p_j(\pi_{kj} - \pi_{hj}) \geq 0 \quad \forall k, h \in A, k \neq h$$

73
To this aim, it suffices to notice that the above condition holds if and only if
\[ \forall k \neq h \neq j \quad \pi_{kk} - \pi_{hk} = \pi_{hh} - \pi_{kh} \text{ and } \pi_{kj} = \pi_{hj} \] (to show this, take \( p^{i=l} = (\frac{1}{2}, \frac{1}{2}, 0, \ldots, 0) \), \( p^{i=l} = (0, 0, \ldots, 1, \ldots, 0) \) respectively). Then, for all \( h \neq k \neq j \):

\[ p_k \geq p_h \Rightarrow \sum_{j=1}^{m} p_j(\pi_{kj} - \pi_{hj}) \geq p_h(\pi_{kk} - \pi_{hk}) + p_h(\pi_{kh} - \pi_{hh}) = 0 \]

\[ \sum_{j=1}^{m} p_j(\pi_{kj} - \pi_{hj}) \geq 0 \Rightarrow p_k(\pi_{kk} - \pi_{hk}) + p_h(\pi_{kh} - \pi_{hh}) \geq 0 \]

\[ \iff \frac{p_k}{p_h} \geq \frac{\pi_{hh} - \pi_{kh}}{\pi_{kk} - \pi_{hk}} = 1 \]

We conclude this Section with a remark. As it is intuitively clear, one could generalize the idea of majority rules to generic \( \alpha \)-majority rules of the kind \( \frac{p_k}{p_h} > \alpha \) and, under appropriate conditions, reproduce exactly optimizing behaviour for an underlying coordination game. However, this would be inconsistent with the motivation we gave in terms of boundedly rational behaviour, in that, any possibility to interpret absorbing states of the dynamics in terms of equilibria, would inevitably require players to know the exact formulation of the rule.

4.4 The Recurrent Communication Classes

We now focus on the characterization of the stochastic process generated under a myopic best-reply with noise at the margin. For reasons entirely analogous to those given when analyzing myopic best-reply in Chapter 3, the process is a (finite) Markov chain. It is clear that, whenever majority rules are payoff maximizing, the dynamics
of a majority rule with noise at the margin have exactly the same absorbing states of a myopic best-reply dynamics of the kind we analyzed in Chapter 3. The notation we use is consistent with that previously introduced to classify the recurrent communication classes of the process under myopic best-reply. We let \( A^{\Gamma_n} (\Gamma) \) and \( A^{\Gamma_b} (\Gamma) \) be the set of absorbing states under majority rules and myopic best-reply respectively.

**Remark 8** Given \( \Gamma = (\Omega, C_m, \mu) \), assume all actions \( i \) and \( l \neq i \) are risk-equivalent. Then \( A^{\Gamma_n} (\Gamma) = A^{\Gamma_b} (\Gamma) \).

We remark that, in a locally interactive model under a majority rule, it is very easy to think of examples in which the dynamics of a majority rule do not converge.

**Example 25** Consider a system of four players on a ‘square’, each having as neighbours the two players not located diagonally opposite them. For any 2-by-2 Coordination game \( A = \{0,1\} \) the following two configurations obviously constitute a 2-cycle of synchronous dynamics, for a majority rule:

\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix} \quad \text{and} \quad 
\begin{bmatrix}
0 & 1 \\
1 & 0
\end{bmatrix}
\]

The convergence properties we are going to derive rely upon the introduction of noise at the margin in the system. In particular, we now relate (the existence of) ergodic sets for a majority rule with noise at the margin to (that of) ergodic sets of an asynchronous myopic best reply dynamics of the kind analyzed in Chapter 3. Notice that the former process "encompasses" the latter: given that every player to whom a change in action is prescribed may or may not follow the prescription, any asynchronous path \( \varphi^n \) can be followed with strictly positive probability by the noise-at-the margin dynamics \( \varphi^n \). Hence, the following holds:
Remark 9 Given \( \Gamma = (\Omega, C_m, \mu) \), assume all actions \( i \) and \( l \neq i \) are risk-equivalent. Then \( E^{p^a}(\Gamma) \neq \emptyset \) \( \Rightarrow \) \( E^{p^a}(\Gamma) \neq \emptyset \).

Proof. Given Remark 8, we know that an entry in \( p^a \) is equal to one if and only if the same holds for \( p^n \). It then suffices to notice that all zero entries of \( p^a \) are also zero entries of \( p^n \). This is enough to prove the claim by standard results. \( \square \)

Given the strict analogy between majority rules and myopic best-reply dynamics for particular classes of coordination games, convergence results are partly based on the findings of which in Section 3.3 of Chapter 3. Recall that, when dealing with uniform local matching models, the results relied on the possibility (allowed by the general formulation) to rule out ties, by an appropriate choice of the size of the neighbourhood. This requirement is implicitly kept in the next result.

Theorem 26 (a) Given \( \Gamma = (\Omega, C_2, \mu_L) \), if \( \Gamma \) has no ties (i.e. it satisfies Assumption 4) and if the two actions are risk-equivalent, then \( A^{p_1}(\Gamma) = E^{p_0}(\Gamma) = A^{p_0}(\Gamma) \).

(b) Given \( \Gamma = (\Omega, C_m, \mu_L) \), if \( \Gamma \) has no ties (i.e. it satisfies Assumption 4), if all actions are risk-equivalent and \( \pi_{ii} = \pi_{jj} \ \forall i \neq j \), then \( A^{p_0}(\Gamma) = E^{p_0}(\Gamma) = A^{p_0}(\Gamma) \).

Proof. Remark 9 shows that cyclic behaviour of the majority-rule-with-noise-at-the-margin dynamics is a sufficient condition in order for the same to hold for the asynchronous myopic best-reply dynamics. To prove the assertion, it is then enough to prove convergence of the latter. This is done by appealing to Theorem 14, part (b), and 13 in Chapter 3 respectively. As for (a), risk-equivalence is needed to guarantee that absorbing states of the majority rule dynamics are the same as those of the myopic best-reply dynamics. As for (b), a further condition is needed in order to guarantee that the underlying game is balanced, as required in Theorem 13\(^2\). Regarding the "no ties" assumption, note that, under the stated assumptions, all the mixed strategy equilibria of the underlying game assign equal probabilities to any

\(^2\)Risk-equivalence does not by itself guarantee symmetry of the payoff matrix.
subset of \( A \), with at least two elements. Provided the population is large enough, a sufficient condition for the assumption to be satisfied is to take \( n \) (the cardinality of the neighbourhood) to be, for example, the first prime number greater than \( m \) (the cardinality of the action space). \( \qed \)

In the section that follows we specialize the model of Section 4.2 to a particular spatial structure. Because we work with a particular type of graph \( \Omega \) we are able to prove convergence for a neighbourhood assignment that does not satisfy the "no ties" assumption.

### 4.5 A Special Structure

The model we study is that of a rectangle with a lattice of horizontally and vertically aligned points which is then folded to form a Torus so that the 'north' boundary of the rectangle is joined with the 'south' boundary, while the 'west' boundary is joined with the 'east' boundary of the rectangle. Each player has as neighbours the eight immediately adjacent players (vertically, horizontally or diagonally) to him on the Torus.\(^3\)

Some of the arguments below use the details of the structure we study in this Section. Diagrammatically, we will represent the structure as a grid of squares. Each player is represented by a square with the neighbouring players being the squares which have at least one edge or vertex in common with it. Thus, one player called 'Centre' (C), with his eight neighbours 'North' (N), 'North-East' (NE), 'East' (E),

---

\(^3\)The results contained in this Section reproduce those in Anderlini and Ianni (1996a).
‘South-East’ (SE), and so on, and his neighbours’ neighbours can be pictured as follows:

<table>
<thead>
<tr>
<th>NW</th>
<th>N</th>
<th>NE</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>C</td>
<td>E</td>
</tr>
<tr>
<td>SW</td>
<td>S</td>
<td>SE</td>
</tr>
</tbody>
</table>

From now on we will refer to the structure we have just described, with \#Ω ≥ 12 players simply as the Ω-Torus (\#Ω is assumed to be even throughout), and we refer to the interaction pattern as µT4. We will take the underlying game to be a 2-by-2 coordination game, usually referred to as C2.

The first thing to notice is that Theorem 26 does not apply to the Ω-Torus, for an underlying C2: if the two equilibria are assumed to be risk-equivalent, then the majority rule does not give player ω a unique prescription whenever actions are chosen with equal frequencies in his neighbourhood.

As in Example 25, it is not difficult to see that the local nature of interaction may prevent convergence of a majority rule dynamics on the Ω-Torus in the absence of noise.

**Example 27** Consider the Ω-Torus with \#Ω = 16. Then the following two configurations clearly constitute a 2-cycle of the system:

---

4The reference is slightly imprecise, in that the set of players is endowed with a specific structure, not reflected in the definition of the interaction pattern we kept throughout this work.
We shall investigate in detail convergency properties of a majority rule with randomization and noise at the margin, as in the following definition:

**Definition 28** Given $C_2$ with action space $A = \{1, 0\}$, $\varphi^n$ is such that for all $\omega$:

\[
P(a_{\theta_i}(\omega) = i \mid a_{\theta_{i-1}}(\omega) = i) = 1 \quad \text{if } \mu_i^{\theta_{i-1}}(\omega) > \frac{1}{2}
\]

\[
0 < P(a_{\theta_i}(\omega) = 1 \mid a_{\theta_{i-1}}(\omega) = i) < 1 \quad l \neq i \quad \text{if } \mu_i^{\theta_{i-1}}(\omega) = \frac{1}{2}
\]

Given the particular spatial arrangement we analyze, we are able to explicitly characterize absorbing states of the process. By a tedious but straightforward case-by-case check, it is in fact possible to show that

**Remark 10** There are two types of absorbing states for the dynamics on the $\Omega$-Torus. The first type is 'homogeneous' in which all players take the same action. The second type is 'mixed' in which some players take one action and other players the other action available in $C_2$. The mixed steady states can only be of one of two forms. The first form is one in which the 'boundary' between areas of the $\Omega$-Torus in which one action is played is always a straight line (vertical or horizontal). In this case the minimum 'thickness' of a set of players playing one particular action is 2. The second form is one in which the boundary between areas of the $\Omega$-Torus in which one action is played is always a '45-degree' line (upward sloping or downward sloping). In this case the minimum 'thickness' of a set of players playing one particular action...
is 3. Diagrammatic examples of the two possible forms of mixed steady states are the following (we ‘unfold’ the entire Ω-Torus as a square)

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

and

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Incidentally, we note that in the above configurations, the aggregate frequencies over unordered couples of players reproduce a correlated equilibrium of the underlying game. For any $C_2$ for which the mixed strategy equilibrium assigns probability $\frac{3}{8} < \alpha < \frac{5}{8}$, the configuration on the left shows an example of an equilibrium of the population game that differs from the replicated version of a Nash-equilibrium of $C_2$: the frequency of action 1 is in fact $\frac{1}{2} < \frac{3}{8}$, but frequencies over pairs of actions do reproduce a correlated equilibrium.

The result that follows proves convergence for the above model, by appealing to a weaker version of Theorem 13 in Chapter 3, together with Remark 9 in this Chapter. Given the neighbourhood assignment, symmetry of the underlying payoff matrix only implies that the average expected payoff in the population, $\Pi(\theta)$, is non-decreasing. The logic of the proof follows two steps: first, it is shown that any ergodic set different from an absorbing state must be contained in the set of states for which the average expected payoff is non-decreasing. Given the symmetry of the matrix, the average expected payoff is non-decreasing.

---

5It can easily be checked that the latter are $[.542, .125]$ and, as such they reproduce a correlated equilibrium of $C_2$.  

80
expected payoff is constant (Lemma 29); second, appealing to specific features of the structure we analyze, it is proved that the latter set is empty.

Lemma 29 Given $\Gamma = (\Omega, C_2, \mu_T)$, assume $C_2$ is such that $\pi_{11} = \pi_{00}$ and $\pi_{01} = \pi_{10}$. Let $L = \{ \theta \in \Theta : \text{there exists no } \theta' \text{ s.t. } P^n(\theta' \mid \theta) > 0, \Pi(\theta') > \Pi(\theta) \}$. Then $E^{\nu^a}(\Gamma) \subseteq L$.

Proof. Given that $C_2$ is symmetric, we know that $\Pi(\theta)$ is increasing whenever a single player switches to his strict best-reply. The latter is not unique, only when a player is surrounded by exactly 4 players choosing 1, in which case any change of action will leave the average expected payoff constant. Hence $\Pi(\theta)$ is non-decreasing along any asynchronous path. Starting from any initial state $\theta_0$, either the system communicates with an absorbing state in a finite number of steps, or it communicates with a state in $L$ (or both). Therefore all (if any) states which do not communicate with an absorbing state, must communicate with a state in $L$. By standard results, this is enough to prove the claim. 

We are now in a position to state and prove our convergence result for the $\Omega$-Torus.

Theorem 30 Given $\Gamma = (\Omega, C_2, \mu_T)$, assume $C_2$ is such that $\pi_{11} = \pi_{00}$ and $\pi_{01} = \pi_{10}$. Then $A^{\nu^a}(\Gamma) = E^{\nu^a}(\Gamma)$.

Proof. From Remark 9 we know that it is enough to show that the asynchronous dynamics admits no non-absorbing ergodic sets or cycles. From Lemma 29, we know that it is then enough to show that $L = \emptyset$. This is the line we will take. We assume, contrarily, that $L$ is not empty.

As a first step we characterize the local configuration of any unstable player in any state $\theta \in L$. It turns out that only two are possible, and this will be crucial later on in the argument. Notice that, by definition of $L$, it must be that, if $\theta \in L$, any player $\omega$ who is not adopting the strict best reply has exactly 4 neighbours choosing 1, since otherwise $\Pi(\theta)$ could be increased by changing $\omega$'s action. It is also easy to see that
the definition of $L$ implies that $\forall \theta \in L$, each $s \in \Omega, \omega$ who adopts the same action as $\omega$ himself, must be stable (otherwise, by changing $s$'s action first and then $\omega$'s action as well, we could increase $\Pi$). With a tedious but straightforward case-by-case check, it is now possible to see that $\forall \theta \in L$ and $\forall \omega$ indifferent between action 1 and action 0, $\omega$'s local configuration must be of the type

\[
\begin{array}{ccc}
1 & 1 & 1 \\
0 & \omega & 1 \\
0 & 0 & 0
\end{array}
\quad \text{or} \quad
\begin{array}{ccc}
1 & 0 & 0 \\
1 & \omega & 1 \\
0 & 0 & 1
\end{array}
\]

We now describe a sequence of transitions starting from a state in $L$ which can take place with positive probability. By the definition of $L$ this implies that the entire path along which we follow the system is in $L$. Let $\theta_0$ be an arbitrary 'initial' state in $L$. Since $\theta_0 \in L$, $\theta_0$ cannot be absorbing. We only deal with the case in which some 1-players are not best-responding in $\theta_0$. The symmetric case in which $\theta_0$ is guaranteed to contain some unstable 0-players is just a re-labelling of this one.

From $\theta_0$, transit to $\theta_1$ by changing to 0 the strategy of exactly one unstable 1-player. Continue in this fashion from $\theta_1$, until a state $\theta_f$ is reached such that either there are no 1-players left, or all 1-players are stable at $\theta_f$. If there are no 1-players left then $\theta_f$ is absorbing, but this is not possible since $\theta_0 \in L$. Therefore the set of 1-players at $\theta_f$ is not empty and contains only stable players.

Choose now an arbitrary unstable 0-player at $\theta_f$. Let this player be denoted by $\omega_f$. Since all 1-players are stable at $\theta_f$ we know that if $\omega$ is a 1-player at $\theta_f$ then $\| \Omega_\omega^1 \| \geq 5$. Since $\omega_f$ is unstable at $\theta_f$, we also know that all 0-players in the neighbourhood of $\omega_f$ are stable at $\theta_f$. In other words, we know that $\omega$ is the only unstable player in the entire neighbourhood.

Transit now to a state $\theta_{f+1}$ by changing only the strategy of $\omega_f$ to 1. Since before the change $\omega$'s local configuration must be as in Figure 6, and all his neighbours must be stable, it is now possible to see with a case by case check (this is simpler if we
notice that the configuration on the right-hand side of Figure 6 is actually impossible in this case) that one of the 0-players in the neighbourhood of $\omega_T$ must be unstable at $\theta_{t+1}$. Let this 0-player be denoted by $\omega_{t+1}$.

Clearly, the number of unstable neighbours of $\omega_{t+1}$ at $\theta_{t+1}$ is exactly one (namely $\omega_t$). This is because all 1-players except $\omega_t$ must be stable and because all 0-players in his neighbourhood must be stable. Transit now from $\theta_{t+1}$ to a state $\theta_{t+2}$ by changing the strategy of $\omega_{t+1}$ to 1. Since $\omega_{t+1}$ has exactly one unstable neighbour at $\theta_{t+1}$, it is possible, again with a case by case check, (we use again the two configurations in Figure 6), and the observation that the configuration on the right-hand side is impossible in this case) to verify that one 0-player in the neighbourhood of $\omega_{t+1}$ must be unstable at $\theta_{t+2}$. Let this 0-player be denoted by $\omega_{t+2}$.

Since $\omega_{t+1} \in \Omega_{\omega_t}$ it must be that $\omega_T$ is now stable at $\theta_{t+2}$ (recall that $\omega_t$ is a 1-player at $\theta_{t+1}$). Therefore, by the same arguments that applied to $\omega_{t+1}$ at $\theta_{t+1}$ we can conclude that exactly one player in the neighbourhood of $\omega_{t+2}$ is unstable at $\theta_{t+2}$ (namely $\omega_{t+1}$).

We can now continue in this way to change to 1 the strategy of unstable 0-players created along the path without bound (note that the number of unstable neighbours of the unstable 0-players along the path is always exactly one). Since all the transitions we have described happen with strictly positive probability, this is a contradiction. If we keep eliminating 0-players, we must eventually reach an absorbing state, but this is impossible since by definition no state in $L$ communicates with an absorbing state in a finite number of steps. Therefore the assumption that $L$ is not empty leads to a contradiction and the proof is complete.  

---

6Recall that we are proceeding by contradiction. Therefore, if all the 0-players in the neighbourhood of $\omega_T$ are stable we already have a contradiction and there is nothing more to prove.
Chapter 5

Aspiration level dynamics: some applications

What it ever comes from sight, hearing, learning from experience:
this I prefer.

(XIV)

5.1 Introduction

In this Chapter we analyze models of population dynamics generated by some specific payoff based learning rules. Rules like myopic best-reply, or majority rules, require players to remember some statistics of the actions played within the neighbourhood in the previous period. Sometimes such knowledge might be difficult, or costly, to collect and retain. In this Chapter we think of players as being less informed than the modeler. The simple contest we analyze is that of a one-shot pure coordination game that agents repeatedly play with their neighbours. In particular, we assume that, at the end of each period, players only remember the payoff obtained in that round.
of interaction. They then compare what they actually got with an aspiration level\textsuperscript{1}. This is a deterministic value, constant across players and over time, obtained as a function of the payoffs of the underlying game. If the payoff exceeds such threshold, they will consider themselves to be satisfied with the action they took and they will not change it in the next period. Instead, if this is not the case, players might decide to do something else, and they may appeal to an underlying learning rule.

Given that players' behaviour is essentially shaped by payoff considerations, the models we analyze in this Chapter depart from previous specifications in that, in every period, each player plays the underlying game at least once. As we shall see, although this constitutes an important conceptual difference, the class of models we analyze has many features in common with those studied in previous Chapters. Analytical results will hence be derived by appealing to the findings of Chapters 3 and 4, applied to an underlying 2-2 Coordination Game. The second part of the Chapter contains some applications.

The Chapter is organized as follows. In Section 5.2.1 we analyze a model where, in each period, players play an underlying 2-2 coordination game with a neighbour chosen at random. We assume that the aspiration level is given by the minimum payoff achievable if coordination is obtained and show that the population dynamics has properties analogous to those studied in Chapter 4. Section 5.2.2 focuses on a model where players play an underlying coordination game with all of their neighbours and hold an aspiration level equal to the payoff associated to the mixed strategy equilibrium of the underlying game. Section 5.3 reports on the results of some simulations obtained with a Cellular Automata Machine. Lastly, in Section 5.4 we interpret the model of Section 5.2.1 as a model of local price search; we find that the local nature of search may be a robust reason for the existence of price dispersion in a search model.

\textsuperscript{1}Different models of aspiration driven learning are, for example, analysed in Bendor, Mookherjee and Ray (1991), Binmore and Samuelson (1995), and Börgers and Sarin (1996).
The model of Section 5.2.1 and the application of Section 5.4 are from Anderlini and Ianni (1996a). A version of the model of which in Section 5.2.2 and part of the simulation results are taken from Anderlini and Ianni (1996b). This Chapter mainly contains some further specifications of the model analyzed in Chapter 4, as well as the application of the results to an explicitly modelled economic situation. The research undertaken in my joint work with L. Anderlini, and later pursued in this dissertation, was triggered by the simulation results summarized in this Chapter. Though simple, the model of local price search constitutes one of the few applications of learning dynamics to an economic model of which I am aware. The substance of both parts of this Chapter warrants future research.

5.2 Aspiration Levels

The general specification we adopt in this Chapter relies on the assumption that players take into account the payoff they got in the last round of interaction in order to decide which action to adopt in the current period. In general, if \( \pi_{t-1}(\omega) \) is the actual payoff to player \( \omega \) at time \( t - 1 \) and \( \hat{\pi} \) is a deterministic threshold quantifying an arbitrary aspiration level, constant across players and over time, the class of rules we analyze is the following:

\[
P(a_{t}(\omega) = i \mid a_{t-1}(\omega) = i) = 1 \quad \text{if } \pi_{t-1}(\omega) \geq \hat{\pi}
\]

\[
P(a_{t}(\omega) = l \mid a_{t-1}(\omega) = i) > 0 \forall l \in A \quad \text{if } \pi_{t-1}(\omega) < \hat{\pi}
\]

As it is clear, the formalization is analogous to the one we adopted in the last Chapter when dealing with majority rules and noise at the margin: players whose payoff falls above the threshold do not change their action between time \( t \) and time \( t - 1 \), and will referred to as stable; players who instead get a payoff that is strictly less than their aspiration level may adopt any available action in the next period.

Further specifications will be described below.

The focus will be on a model of uniform local matching, as in Definition 10, for an underlying 2-2 Coordination Game, for which actions are labelled 1 and 0.

Given that players base their choices on the payoff of the last interaction, it becomes essential to adapt the model we have used so far, to account for a situation where all players play the underlying game at least once in each time period. Recall that the interaction pattern \( \mu \) is defined by a probability distribution over a subset of all unordered pairs of players. We would like to specify conditions under which \( \mu \) produces a complete matching among players. By that we identify a specific coupling, ensuring that each player in the population is matched with one (and only one) of his potential opponents. Once again, if we represent the set of players, \( \Omega \), as a graph (as in Section 4.2 of Chapter 4) things are easier, in that the requirement is formalized by the following assumption:

**Assumption 5** \( \Omega \) admits at least one decomposition into the product of \( n \) 1-factors.

Assumption 5 is satisfied in all the specific spatial arrangements we have described so far; it is however not true in general\(^2\). From \( \mu \), we then define a probability distribution over the set of \( \Omega \)'s 1-factors, and think of one round of interaction as being determined by a random draw of one, out of \( n \), 1-factor. Throughout the work we shall assume that matching is uniform; as a direct implication, the probability distribution over the set of 1-factors is also uniform. The actual payoff will then be

\[^2\text{It is known (Andrásfai (1977)) that some finite } n \text{-regular graphs (with an even number of vertices) do not possess any 1-factor. Sufficient conditions for this to hold are known, but necessary and sufficient conditions for a finite } n \text{-regular graph to have any 1-factor are quite complex indeed (Bollobás (1979)). If the graph is bipartite, than a necessary and sufficient condition for the existence of a complete matching is known as Hall's condition (Biggs (1989)). We conjecture that an alternative way of proceeding might be possible. One could achieve a coherent random coupling pattern starting locally with one random match, and then continuing 'outward', and sequentially across the system with random matches which are constrained not to involve any players which have already been coupled with one of their potential opponents. Problems related to this way of proceeding are analyzed in Follmer (1974) and Kirman, Oddou and Weber (1986).}\]
specified accordingly. As for the specific value of the aspiration level, we think of it as a simple function of the payoffs of the underlying game.

5.2.1 Majority Rules and Payoff memory

In this Subsection we specify a model where all players play an underlying game exactly once in each time period. The model is based on the idea that players take the realization of their random payoff at $t-1$ into account when deciding how to play at $t$. Imagine that the system is in a given state $\theta$ and player $\omega$ has been randomly coupled with one of his neighbours $s \in \Omega_\omega$ to play an underlying $C_2$. $\omega$ then receives a payoff which is given by:

$$\pi(a_\theta(\omega), a_\theta(s))$$

Let $\tilde{\pi}_\theta(\omega)$ represent $\omega$'s random payoff when the state of the system is $\theta$. Then, under Assumption 5 and if matching is uniform, the probability distribution of $\tilde{\pi}_\theta(\omega)$ is:

$$\tilde{\pi}_\theta(\omega) = \begin{cases} 
\pi_{11} & \text{with probability } \frac{\#\Omega_\omega^1}{n} \\
\pi_{10} & \text{with probability } 1 - \frac{\#\Omega_\omega^1}{n} \\
\pi_{00} & \text{with probability } 1 - \frac{\#\Omega_\omega^0}{n} \\
\pi_{01} & \text{with probability } \frac{\#\Omega_\omega^0}{n}
\end{cases}$$  \hspace{1cm} (5.1)

where $\pi_{ij}$ are the payoffs of $C$, $\#\Omega_\omega^i$ is the number of $\omega$'s neighbours playing $i$, and $n = \#\Omega_\omega$.

We now turn to the specification of the aspiration level. Given the nature of the underlying game, it seems intuitively appealing to say that $\pi_{11}$ and $\pi_{00}$ are the 'good' payoffs whereas $\pi_{10}$ and $\pi_{01}$ are the 'bad' payoffs of $C$. This intuition is correct only if $C_2$ also satisfies the following assumption:

Assumption 6 $C_2$ is such that $\pi_{11} > \pi_{10}$ and $\pi_{00} > \pi_{01}$.  

89
The assumption implies that, for each action, the expected payoff is strictly increasing in the probability with which the opponent chooses the same action. If it is satisfied, then we can think of the aspiration level as being given by the minimum payoff achievable if coordination is obtained, i.e. $\hat{\pi} = \min\{\pi_{00}, \pi_{11}\}$. We then formulate the following behavioural rule:

$$P(a_{\theta_1}(\omega) = i \mid a_{\theta_{t-1}}(\omega) = i) = 1 \quad \text{if } \hat{\pi}_{\theta_{t-1}}(\omega) \geq \hat{\pi}$$

$$P(a_{\theta_1}(\omega) = l \mid a_{\theta_{t-1}}(\omega) = i) > 0 \quad \forall l \in A \quad \text{if } \hat{\pi}_{\theta_{t-1}}(\omega) < \hat{\pi}$$

Assumption 6 guarantees that a player who is adopting an equilibrium action, will necessarily achieve his aspiration level and hence will not change his action in the next period. As an obvious implication, a state where all players adopt exactly the same action is an absorbing state of the dynamics. The same does not apply to absorbing states where different actions are adopted in the population. To see this, consider an absorbing state of the kind depicted in Figure ??, of Section 4.5 in Chapter 4: if a player adopting action 1 is matched with a player adopting action 0 (and this will happen with probability $\frac{1}{2}$), then he will not achieve his aspiration level. Hence, so the rule postulates, he may adopt any of the two available actions in the following period. As a result, such configuration cannot be stable. In essence, the reason is that the rule, in the above formulation, does not allow players to hold different "views of the world".

A different specification of the above rule, that allows for different actions to coexist in an equilibrium of the population game is based on an underlying majority rule, analogous to those we considered in Chapter 4. The underlying idea is that whenever the payoff falls below the aspiration level, a player adopts a majority rule. We call this rule a majority rule with payoff memory, since it is the payoff in the last period which determines the players' behaviour together with the local configuration of the system, and we specify it as follows:

**Definition 31** The transition matrix for a Majority rule with payoff memory is de-
noted as $p^q$. Each entry is such that for all $\omega$:

$$
\begin{align*}
P(a_{\theta_i}(\omega) = i | a_{\theta_{i-1}}(\omega) = i) &= 1 & \text{if } \bar{\pi}_{\theta_{i-1}}(\omega) \geq \hat{\pi} \\
a_{\theta_i}(\omega) &= 1 & \text{if } \#\Omega^1_{\omega}/n \geq 1/2 \\
a_{\theta_i}(\omega) &= 0 & \text{if } \#\Omega^1_{\omega}/n < 1/2
\end{align*}
$$

The rule postulates that a player who achieves his aspiration level, will stick to the action previously adopted, while a player whose payoff falls below the threshold will adopt a majority rule. As it can easily be seen, noise at the margin and payoff memory have similar effects on majority rules: a majority rule and payoff memory is locally equivalent to the same majority rule with noise at the margin of a variable degree, which depends on $\theta$, on the player $\omega$ and on the configuration of play in $\omega$'s neighbourhood.

We are going to prove that the system described by $\varphi^p$ converges to an absorbing state in finite time. We do so by appealing to Theorem 14 of Chapter 3. Two caveats apply to the similarity between noise at the margin and payoff memory\(^3\). The first thing we must be careful about is the possibility of isolated players. If at a particular state a player is 'surrounded' by players who are playing a strategy different from his own, then the probability that player will achieve his aspiration payoff level is zero. This means that he will switch strategy with probability one in the next period as opposed to switching with a probability which is positive but strictly less than one in the noise at the margin case. The second difficulty comes from the fact that, a player will achieve his aspiration level payoff if and only if the neighbouring player with whom he has been coupled also achieves his aspiration level payoff. This creates a degree of correlation in the noise induced by payoff memory which is not present in the case of noise at the margin. The delicate case turns out to be the one in which two players playing the same strategy are surrounded entirely by players who

---

\(^3\)This model is analog in content to the one studied in Anderlini and Ianni (1996a). The authors are grateful to an anonymous referee for specific comments on the 'equivalence' between noise at the margin and pay-off memory.
play a different strategy. In this case with payoff memory, either both players will change strategy in the next period or neither of them will. In the case of noise at the margin the same situation yields no switches, one switch and two switches, all with strictly positive probability. Throughout, we will refer to pairs of players whose local configuration is as above in a particular state $\theta$ as ‘isolated pairs’ at $\theta'$.

The two-phase construction we have used to prove the second statement of Theorem 14 relies on the flexibility in the number of unstable players which change action in the period following any unstable state of the system, guaranteed by noise at the margin. Close inspection of the construction, reveals that all that is needed is that with strictly positive probability none of the unstable players playing a particular strategy change their action in the following period. Clearly, for these two features to hold we need to exclude states with isolated players, while states with isolated pairs present no problem. Inspection of Definition 31 is enough to prove the following.

**Remark 11** Consider a majority rule with payoff memory. Fix a strategy $i \in \{0; 1\}$. Let $\theta$ be an unstable state such that some $i$-players are unstable at $\theta$. Then we have that a) the system transits with positive probability to a state $\theta'$ in which all unstable $i$-players change strategy, and b) if there are no isolated $i$-players, the system transits with strictly positive probability to a state $\theta''$ in which none of the unstable $i$-players change strategy.

We are now ready to prove that the system converges to an absorbing state, and that absorbing states under a majority rule with payoff memory ($E^p$) are the same as the absorbing states under a majority rule with noise at the margin ($E^n$):

**Theorem 32** Given $\Gamma = (\Omega, C_2, \mu_L)$, assume that the two equilibria are risk-equivalent and that payoffs satisfy Assumption 6. Then $A^p(\Gamma) = E^p(\Gamma) = A^n(\Gamma)$.

**Proof.** As with Theorem 14, we only need to show that all states of the system ‘communicate’ with some absorbing state with strictly positive probability in a finite
number of steps. Consider an arbitrary initial unstable state $\theta_0$. We only deal with the case in which some 1-players are unstable at $\theta_0$, the other case is only a re-labelling of this one. Two cases are possible. Either the set of isolated 0-players at $\theta_0$ is empty or it is not. If it is empty, then the proof of Theorem 14 applies exactly as it is. If $\theta_0$ contains some isolated 0-players, first transit to $\theta_1$ changing the strategy of all unstable 0-players. By Remark 11, it is then sufficient to show that no isolated players will appear along the transitions. But this is indeed the case, because, at each step, the player who changes action conforms to a majority of players. Hence, he cannot be unstable. Since we start the procedure at $\theta_1$, after having changed the strategy of all isolated 0-players (if any), this is enough to prove the claim. 

For the sake of completeness we also state the convergence result for a majority rule with payoff memory and randomization (i.e. when players are allowed to adopt a mixed strategy if $\|\Omega^1/n = 1/2$), when players are located on the $\Omega$-torus (the proof of Section 4.5 in Chapter 4 applies unaltered).

Theorem 33 Given $\Gamma = (\Omega, C_2, \mu_T)$, assume that the two equilibria are risk-equivalent and that payoffs satisfy Assumption 6. Then $A_{p\mu}(\Gamma) = E_{p\mu}(\Gamma) = A_{p\mu}(\Gamma)$.

5.2.2 Aspiration Levels

The model we analyze in this Section is just a slight modification of the one of the previous Section. The main departure from the model we described in the previous Section is that, in each round of interaction, players play the underlying 2-2 Coordination game with all their neighbours. In particular, we assume that, at the beginning of period $t$, after having played $C_2$ with all his neighbours in the previous period, player $\omega$ remembers only the average payoff he obtained in that round of interaction. On the basis of the latter he decides which action is to be adopted; specifically, he compares the average payoff he obtained with a threshold, constant across players and over time, that defines his aspiration level. The average payoff for player $\omega$ at $t$ is obviously a function of his action and the actions taken within his neighbourhood; we
denote it by: \( \pi_{\theta_{t-1}}(\omega) \), and we obtain it as the expected value of the random payoff of which in (5.1). We then assume that the aspiration level is given by the payoff corresponding to the (unique) mixed strategy equilibrium of \( C_2 \), that we denote by \( \hat{\pi} \). We postulate the following \textit{aspiration level} updating rule:

\textbf{Definition 34} The transition matrix for an Aspiration Level updating rule is denoted as \( \varphi^{asp} \). Each entry is such that for all \( \omega \):

\[
P(a_{\theta_t}(\omega) = i \mid a_{\theta_{t-1}}(\omega) = i) = 1 \quad \text{if} \quad \pi_{\theta_{t-1}}(\omega) \geq \hat{\pi} \\
P(a_{\theta_t}(\omega) = l \mid a_{\theta_{t-1}}(\omega) = i) > 0 \quad \forall l \in A \quad \text{if} \quad \pi_{\theta_{t-1}}(\omega) < \hat{\pi}
\]

where \( \hat{\pi} \) is the payoff associated to the mixed strategy Nash-equilibrium of \( C_2 \).

If the payoffs of \( C_2 \) satisfy Assumption 6, i.e. if the expected payoff associated to each action is increasing in the probability with which a potential opponent chooses the same action, then the payoff achievable in the mixed strategy equilibrium is exactly the minimum payoff achievable by adopting the action that maximizes the expected payoff. As a result, by the mere definition, players who adopt the above rule are adopting the best-reply in any stationary state\(^4\). Furthermore, in a model of uniform local matching, given that each player plays only with his neighbours, and these are a subset of players in the population, players may hold, in equilibrium different views of the world, and, accordingly, adopt different actions. The specification of the behaviour of players when the average payoff obtained falls below the aspiration level, allows for the flexibility required in the proof of the convergence result of Theorem 26 in Chapter 4, that, for this model applies unaltered. In words, the next Theorem, says that if players adopt the above rule when playing an underlying 2-2 coordination game for which Assumption 6 holds, then a) the system will converge to an absorbing state in finite time and b) absorbing states of the dynamics \( \varphi^{asp} \) coincide with the set

\(^4\)As it is perhaps obvious, Assumption 6 is needed to justify the only if part of the equivalence between best-reply behaviour and the aspiration level learning rule we formulate.
of absorbing states under asynchronous myopic best-reply \( \varphi^a \). We state the result for a model of uniform local matching.

**Theorem 35** Given \( \Gamma = (\Omega, C_2, \mu_L) \), assume that payoffs satisfy Assumption 6. Then \( A^{a_{\text{ap}}} (\Gamma) = E^{a_{\text{ap}}} (\Gamma) = A^a (\Gamma) \).

### 5.3 Simulation Results

To characterize the steady-states of the system, we rely on the results of simulations which we carried out using a "Cellular Automaton Machine"\(^5\). The prototype of what is termed a Cellular Automaton is informally described by the following set-up. Consider a finite chessboard in which each square can be of one, out of \( m \), colours at each moment in time. Assume that we have a rule that specifies what the colour of each square should be as a function of its four neighbouring squares — North, South, East and West. Now let an initial pattern of coloured squares be given at time \( t = 0 \). We then turn the system on and let the rule of state transition operate, examining the pattern that emerges as \( t \to \infty \). In other words, we watch the squares taking different colours at each time step in accordance, and look for the steady-state pattern, if any, that emerges as \( t \) becomes large. It is intuitively clear that the key ingredients of the set-up are exactly those we are using in a model of local matching. In particular, our requirements of *homogeneity* and *locality* are met: each cell of the system is the same as any other (in the sense that cells can each take on exactly the same set of \( m \) colours), and the transition rule is local both in space and time (in that we do not have any time-lag effects, nor do we have non-local interaction affecting the state).

More accurately, this setting describes a two-dimensional cellular automaton, since

\(^5\)Casti (1991) provides an excellent introduction to the dynamics of Cellular Automata, as well as an exhaustive bibliography. As further references, we recall the work of Toffoli and Margolus (1987) and Wuinsche and Lesser (1992). The simulations have been carried out on a CAM-PC hardware board. The programmes we used are available on request. CAM-PC is © of AUTOMATRIX, INC., P.O. Box 196, Rexford, NY 12148-0196 U.S.A.
the state-space is a planar grid. There are, traditionally, two basic neighbourhoods of interest: the Von Neumann neighbourhoods (those of the example above), and the Moore neighbourhoods (which also include those cells that are diagonally adjacent).

We are now ready to describe more in detail the set-up of our simulations. We assume players live on a 2-dimensional lattice of $256 \times 256$ cells, each cell constituting one player's address. Each player interacts repeatedly only with his eight closest neighbours, i.e. North, North-East, East, South-East, South, South-West, West and North-West. Hence the cardinality of the neighbourhood is fixed and equal to 8. We chose to deal with Moore-neighbourhoods because this, loosely speaking, guarantees symmetry in any direction. Moreover, instead of making explicit provisions for dealing with boundaries (where players would inevitably be surrounded by less than 8 neighbours), we assume that the boundary is periodic: precisely, we assume that the square grid is folded to form a Torus.

When a player interacts with one of his neighbours, he plays the one-shot normal form coordination game, $C_2$. We take the underlying game to be:

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

More precisely, for any underlying 2-by-2 coordination game, the behavioural assumption we postulate, yields the same dynamics of a totalistic\textsuperscript{6} rule applied to $C$. To see this, identify the address of player $\omega$ by the coordinates $k$ and $l$. Assume Moore neighbourhoods so that

\[
\forall \omega \quad \Omega_\omega = \{(k, l + 1), (k, l - 1), (k + 1, l), (k - 1, l), (k + 1, l + 1), (k - 1, l - 1), (k + 1, l - 1), (k - 1, l + 1)\}
\]

Recall that a 2-2 coordination game has a unique mixed strategy equilibrium. Let

\textsuperscript{6}Totalistic rules are those where the value of a cell at time $t$ depends only on the sum of the values of the cells in the neighbourhood at time $t - 1$.}

96
the latter assign probability $\alpha$ to action 1. Then, an equivalent formulation of best-reply dynamics applied to a 2-2 coordination game, can be obtained by "adding up" all the neighbours choosing action 1 and defining a threshold $m = 8\alpha$. We can hence specify $\omega$'s optimal behaviour according to whether the sum of 1-players in $\omega$'s neighbourhood, $\sum_{s \in N_\omega} a(s) = \|N_\omega\|$, is above, below or equal to $m$. Best-reply behaviour would prescribe player $\omega$ to adopt action 1, action 0 or to randomize, respectively. By rescaling the cut-off value $m$, we can therefore consider any generic 2-2 coordination game.

We started by considering $m = 4$, that is the case where the two Nash equilibria in pure strategies are risk-equivalent. The dynamics of the simulations are not difficult to describe in words. Consider a small cluster of players playing one strategy; the players on the 'edges' are necessarily unstable, i.e. not surrounded by a majority of players choosing the same action they do, and, therefore, likely to switch. The small cluster will thus be progressively 'eroded' and eventually disappear. Consistently with Section 4.5 of Chapter 4, in any stable configuration, players must be aligned on a straight-line in a 'stripe' wrapped around the Torus. The thickness of such stripes must be of at least two players if horizontal or vertical, and at least three players if diagonal. Any discontinuity will generate a process of erosion of the cluster that will inevitably lead to its disappearance.

Things become relatively more complicated if $m \neq 4$. Here as well the system could indeed converge to a stable state in which both strategies coexist, but the patterns that emerge are more difficult to be described. Relying on imagination, rather than on formalism, we could describe the stable clusters that emerge in the long run as 'spots', the maximum curvature of which is entirely determined by the cut-off value. In general, one weak property of the self-organizing pattern that emerges from the simulations, is worth being noted: 'stripes' and 'spots' cannot coexist and as the difference in terms of risk between the two equilibria becomes less important (i.e. as $m \to 4$) the spot pattern connects up into a pattern of stripes.
We conclude with a question which naturally arises from this work and warrants future attention. In short, one would like to know 'how likely' are the mixed steady states in which two strategies coexist? The question can be divided into two more specific ones. The first is, given a particular initial distribution of strategies across the spatial structure, how likely is the convergence to a mixed steady state? The second step is to ask how wide is the range of initial distributions of strategies which allows convergence to a mixed steady-state with positive probability, or more generally how the likelihood of mixed steady states changes as we vary the initial distribution of actions.

We do not have analytic answers, but can tentatively report some patterns which emerged during the simulations. Our simulations all started with a random draw of strategy for each player with equal probability, and seem to give convergence to mixed steady states in roughly one fifth of the cases (the vertical and horizontal boundaries being roughly twice as likely as the diagonal ones). It should be strongly emphasized that these patterns are in our view to be treated with extreme caution because of substantial programming short-cuts we were forced to use. In the first place, the hardware we used made it impossible to obtain a proper coupling of players across the Torus. In essence in our simulations each player played one of his neighbours randomly, regardless of whether that player had been coupled with another player or not. Moreover, for a structure with the order of 65,000 players it is extremely difficult to generate 'enough noise' at sufficiently high speed. In order to introduce the noise at the margin in the behaviour of players, our simulations relied on 'pseudo-noise' generated through a highly non-linear rule unrelated to the system to be simulated.

5.4 Shopping on a Torus

In this Section we present an interpretation of the model of Section 5.2.1 as a model of local price search. We take the $\Omega$-Torus case purely for convenience. The numerical
values of the parameters of the analysis of this Section could be changed in order to fit other spatial arrangements to which the results apply.

Consider the simplest search model with identical buyers of a single homogeneous good. Assume we have a finite population of \(\Omega\) sellers, each of them representing a shop, located at a site on the \(\Omega\)-Torus. Customers wish to purchase exactly one unit of the commodity. At the beginning of every time period, \(t\), one buyer arrives at each shop. Shops commit themselves to a price at which they must serve all the clients they get. Shops can choose between two strategies; charging a high price, say \(P_H\), or a low price, \(P_L < P_H\). Furthermore, we assume that all shops face the same constant marginal cost \(c < P_L\) and have no fixed costs.

Shops seek to maximize their profit facing the uncertainty that derives from the behaviour of customers: a buyer faced with a high price, may walk out and search for a lower price, but only within the neighbourhood of the shop to which he is originally assigned. A buyer who searches observes the price of one of the neighbouring shops with a uniform probability of \(1/8\). Search is costly. A buyer who engages in search incurs a search cost, identical for all buyers, equal to \(q\). The probability of finding a lower price, obviously depends on the number of neighbouring shops charging \(P_L\).

We assume that the search cost is sufficiently small so that the expectation that at least one of the neighbouring shops charges a low price is sufficient to induce customers who observe a high price to search. In other words:

**Assumption 7** The search cost \(q\) satisfies

\[
0 < q < \frac{1}{8}(P_H - P_L)
\]

We assume that the buyers correctly perceive the probability that search will result in the observation of a low price as being equal to the fraction of neighbouring shops actually charging \(P_L\).\(^7\) The behaviour of customers can therefore be summarized as

\[^7\text{It may seem that we are endowing our shoppers with excessive 'rationality' for a world in which}\]
follows: a) a buyer faced with a low price, buys from the shop he is initially assigned to and b) a buyer faced with a high price buys from the shop he is assigned to, only if all the neighbouring shops charge $P_H$, otherwise he always pays the search cost $q$ and therefore observes the price of one of the neighbouring shops drawn at random; if the neighbour drawn at random charges strictly less than the price initially observed, then the customer moves and buys from the cheaper shop.

Under these assumptions, by choosing the high-price strategy, $P_H$, the shop will get at most one customer, whereas by charging the low price, $P_L$, it will sell at least one unit of product. It is evident that the expected payoff associated with a price strategy, depends on the strategies adopted within the neighbourhood. Intuitively, it is clear how a trade-off exists between selling more units at a low price and fewer units at a higher price. Since the maximum number of units that any shop can possibly sell is limited to the number of neighbouring shops (eight), the difference between the high price and the low price must not be too high to make the model interesting. Formally we assume:

**Assumption 8** The prices $P_H$ and $P_L$ are such that

$$\frac{1}{3}P_H < P_L < P_H$$

Let $h(\omega; \theta)$ be the number of $\omega$'s neighbours charging $P_H$ when the system is in state $\theta$. Straightforward algebra shows that the expected payoffs, $\pi_H^E(\omega; \theta)$ and $\pi_L^E(\omega; \theta)$, associated with the two price strategies, $P_H$ and $P_L$ are as follows

$$\pi_H^E(\omega; \theta) = (P_H - c)\frac{h(\omega; \theta)}{8} \quad \text{and} \quad \pi_L^E(\omega; \theta) = (P_L - c)(1 - \frac{h(\omega; \theta)}{8})$$

---

*behaviour of other agents is dictated by myopic learning rules. We do this since it makes the steady states of the dynamics we study in this section correspond to those we described before. The analysis we carry out here would, however remain unchanged if we assumed that the buyers have some myopic probability of successful search in mind, and the search cost is sufficiently low so as to ensure that all buyers who observe a high price engage in the costly search.*
Given the values of $P_H$ and $P_L$, the precise point of balance between the two strategies depends on the marginal cost $c$. Therefore, by an appropriate choice of $c$ we can reproduce a majority rule behaviour like the one described in Section 31. Formally we state

**Remark 12** If (and only if)

$$c = \frac{(3P_L - P_H)}{2}$$

then the expected payoffs associated to the two price strategies, $P_H$ and $P_L$, are such that:

$$\pi^E_H(\omega; \theta) \geq \pi^E_L(\omega; \theta) \iff h(\omega; \theta) \geq 4$$

It is useful to keep track of the actual probability distributions of profits and of profits per unit sold associated with the two pricing strategies.

**Remark 13** The profits per unit sold when $c = (3P_L - P_H)/2$ are $(3/2)(P_H - P_L)$ for the shops charging $P_H$ and $(1/2)(P_H - P_L)$ for the shops charging $P_L$. The probability distribution over units sold for shops charging $P_H$ is

- 1 with probability $h(\omega; \theta)/8$
- 0 with probability $1 - h(\omega; \theta)/8$

and for the shops charging $P_L$:

$$(1 + k) \text{ with probability } \left( \frac{h(\omega; \theta)}{k} \right) \left( \frac{1}{8} \right)^k \left( \frac{7}{8} \right)^{h(\omega; \theta) - k} \quad k = 0; \ldots; h(\omega; \theta)$$

We denote the random payoff to shop $\omega$ in state $\theta$ by $\tilde{\pi}(\omega; \theta)$.

---

\(^{8}\text{Note that if } h(\omega; \theta) = 0 \text{ the formula below implies that the shop sells one unit with probability one.}\)
The dynamics we are interested in are the ones which arise from the shops using the majority rule implicit in Remark 12, with payoff memory. In other words, if a shop's profit does not fall below a given aspiration level which we choose to be equal to \( P_H - P_L \), then the shop will simply carry on charging the same price without change. If on the other hand a shop's profit falls below \( P_H - P_L \), then the price charged in the next period will be determined by a \( 4-4 \) majority rule consistent with myopic profit-maximization. We also assume that when a shop is indifferent in expected terms between charging \( P_H \) and \( P_L \), the majority rule involves randomization. Formally we have

**Definition 36** The majority learning rule with randomization and payoff memory for the shopping model on the \( \Omega \)-Torus is such that

\[
\begin{align*}
P(a_{\theta_i}(\omega) = i \mid a_{\theta_{i-1}}(\omega) = i) &= 1 & \text{if } \bar{\pi}(\omega; \theta) \geq P_H - P_L \\
a_{\theta_i}(\omega) &= P_H & \text{if } h(\omega; \theta) > 4 \\
a_{\theta_i}(\omega) &= P_L & \text{if } h(\omega; \theta) < 4 & \text{if } \bar{\pi}(\omega; \theta) < P_H - P_L \\
P(a_{\theta_i}(\omega) = P_H) = P(a_{\theta_i}(\omega) = P_L) & \text{ if } h(\omega; \theta) = 4
\end{align*}
\]

It is interesting to see precisely how the dynamics are specified. Let us start with a shop charging \( P_H \). If more than four neighbouring shops also charge \( P_H \), two cases are possible. Either the shop will sell one unit or none (the latter is only possible if some neighbouring shop charges \( P_L \) and the customer's search reveals this). If one unit is sold at \( P_H \) the realized profit is more than \( P_H - P_L \) and therefore the shop will keep charging \( P_H \) in the next period. If no units are sold then the profit is zero, but since more than four neighbours charge \( P_H \) the shop will keep charging \( P_H \) in the next period. If exactly four neighbours charge \( P_H \), again either one unit is sold or no units are sold. If one unit is sold, the shop will keep charging \( P_H \) in the next period. If no units are sold the shop's profit is below the aspiration level \( P_H - P_L \), and therefore the fact that exactly four neighbours charge \( P_H \) will induce the shop
to randomize between $P_H$ and $P_L$ with equal probability in the next period. The key feature of this case is that overall the shop will change to $P_L$ and will stick to $P_H$ both with strictly positive probability. If a shop charging $P_H$ faces a majority of shops charging $P_L$ but is not isolated, then again it will sell either one or zero units, both with strictly positive probability. Therefore it will change to $P_L$ and stick to $P_H$ both with strictly positive probability. A shop charging $P_H$ which is isolated (surrounded by shops charging $P_L$) will lose its customer with probability one and therefore change to $P_L$ with certainty in the next period. Notice that the behaviour of isolated pairs of shops surrounded by shops charging $P_L$ is not perfectly correlated as in the case of payoff memory of Section 31. This is because it is possible that one shop loses its customer to a cheaper shop but the other does not because the buyer’s search reveals the only other high price in the neighbourhood.

Consider now shops charging the low price $P_L$. Observe that shops charging $P_L$ achieve their aspiration profit level of $P_H - P_L$ only if they sell two or more units. Take the case of a shop charging $P_L$ with more than four neighbours charging $P_L$. It may sell one or more units (the latter is only possible one or more neighbours charge $P_H$). If two or more units are sold the aspiration level of profit is achieved and therefore the shop will keep charging $P_L$ in the next period. If only one unit is sold (and it is interesting to notice that this will certainly be the case if all the shop’s neighbours charge $P_L$) the profit is below $P_H - P_L$, but since more than four shops in the neighbourhood charge $P_L$ the shop will keep charging $P_L$ in the next period. If precisely four neighbours of a shop charging $P_L$ also charge $P_L$, then again one or more units may be sold. If only one unit is sold the shop will randomize between $P_H$ and $P_L$ with equal probability in the next period. If two or more units are sold the shop will keep charging $P_L$. Overall, the shop will change price and stay put both with positive probability. If more than four neighbours of a shop charging $P_L$ charge the high price $P_H$, the shop will change to $P_H$ with certainty if only one unit is sold and will stick to $P_L$ if two or more units are sold. Notice that one and two or
more units will be sold all with strictly positive probability in this case. This is true even in the case in which a shop charging $P_L$ is isolated in the sense of having eight neighbours charging $P_H$. Therefore isolated shops charging $P_L$ only change strategy with probability less than one. This is in contrast with the payoff memory rule of Section 31. Lastly, note that the behaviour of isolated pairs of shops charging $P_L$ is not perfectly correlated since one of them may get additional customers while the other does not.

From the discussion above it is clear that the dynamics given by are qualitatively the same as those of Section 31, with the exception of isolated pairs of both pricing strategies and of isolated shops charging $P_L$. A formal modification of the argument is not strictly needed, however. This is because under isolated pairs of players change strategy simultaneously with strictly positive probability, and all isolated players change strategy with strictly positive probability. The proof of Theorem 32 applies unchanged to the following result.

**Corollary 37** Consider the Markov chain given by the majority rule with randomization and payoff memory for the shopping model on the $\Omega$-Torus. Starting from any initial state, the system will converge to an absorbing state in finite time with probability one.

Once convergence is established, the equilibrium configuration are of particular interest. Locally interactive systems may provide a robust justification for price dispersion in a search model.

**Remark 14** Absorbing states for the shopping model with local search on the $\Omega$-Torus may be 'mixed'. This amounts to say that, in the search model described above, the system might converge to an equilibrium configuration in which some shops charge $P_H$ and some others charge $P_L$.

In other words, starting from a situation of complete ex-ante homogeneity (in costs, tastes and behaviour), locally interactive systems, may give rise to heteroge-
neous pricing behaviour. The characterization of such possible ‘mixed’ steady states is entirely determined by the specific spatial structure assumed.

The possibility of equilibrium price dispersion in a search model is not a new result in the literature. However, equilibrium price dispersion has often been driven by some sort of heterogeneity either in cost of production (Reinganum (1979), MacMinn (1982)), or in the search cost that consumers have to pay (Salop and Stiglitz (1976, 1982)), or in the propensity to search itself (Schwartz and Wilde (1979)). In contrast, we obtain equilibrium price dispersion in a model in which neither costs, nor tastes, differ across agents. Our results also differ from Burdett and Judd (1983) in that we do not need to assume that search is noisy, in the sense that one search shot yields, with positive probability, more than one price observation for the consumer.

The formalization presented in this Section allows for a quite general class of economic applications, ranging from mimetic contagion in financial markets (as in Kirman (1991)), to dynamics of technological diffusion (as those studied in a non-local setting in Dosi and Kaniovski (1993)). In general, there seems to be no a priori reason not to extend the analysis to many economic domains, in which the following key ingredients are present: existence of (positive) network externalities, complex communication structure in the population that motivates boundedly rational ‘imitative’ behaviour on the part of agents. As the formalization suggests, while path-dependency applies throughout, informational imperfections and/or the complexity of the environment, fosters variety, i.e. the coexistence, in equilibrium, of different strategy choices.
Chapter 6

Myopic Best-Reply with Mistakes

It rests by changing.
(LII)

6.1 Introduction

In this Chapter we complement the analysis of Chapter 3 by focusing on some models of myopic best-reply dynamics with mistakes. Mistakes in the decision process undertaken by players, model perturbations that affect the underlying dynamics. Such perturbation is customarily referred to as the noise. Extensively analyzed in the recent literature, noisy dynamics formalize the idea that a small amount of random noise in players' actions may drive population behaviour towards particular equilibria and away from others. As a result, some equilibria may be selected by the dynamics as the most likely outcome of play when interaction is repeated over time.

A generic model of noisy best-reply dynamics involves at least two aspects: the specification of the noisy decision process on the part of players and the study of the dynamics of the aggregate population. As far as the first is concerned, we take the underlying noiseless process to be exactly the myopic best-reply dynamics of Chapter 3. However, the latter is "perturbed", in that we assume that when players are called
to choose actions, they best-respond with high probability, but, with some small probability, they do something else instead.

We shall focus on two different specifications of a model of population dynamics. In the first, *noisy best-reply dynamics*, we assume that each player makes a mistake with a probability that is fixed, equal for all players and uncorrelated over time. As we shall see, the dynamics of the aggregate population is described by a Markov chain over the same state space of the underlying myopic best-reply dynamics. Besides, the latter is *regular*, in that all transitions may occur with strictly positive probability. Regular Markov chains are known to admit a unique limit distribution (in the terminology of Chapter 3, this means that the process admits a single ergodic set, that contains *all* states). Uniqueness of the limit distribution then motivates further characterizations of the process as the "noise" becomes negligible. These dynamics have been studied extensively in the recent literature on equilibrium selection; in the first part of this Chapter, we report the main findings and focus on particular cases that seem to undermine the generality of the results.

In the second part of this Chapter, we analyze a model where the probability with which mistakes occur explicitly depends on the expected payoff from the interaction. We refer to this model as *best-reply dynamics with payoff dependent mistakes*. Whenever expected payoffs differ across players, then so do the probabilities with which each player adopts each single action. This introduces a high degree of correlation among players' choices. As a result, though the process still inherits Markovian properties from the underlying myopic best reply dynamics, aggregate transition probabilities depend on the specific configuration of play. We shall formalize the process as a Markov random field. For some classes of games we are able to characterize it in terms of the set of Gibbs measures associated to a particular specification of the interaction potential, given by the average expected payoff in the population, introduced in Chapter 3. The behaviour of this category of processes is, in general, less well understood than that of simple Markov chains. Under asynchronous dynamics,
we are able to explicitly derive the unique limit distribution and, in analogy with what we do in the first part of this Chapter, address equilibrium selection issues.

The analysis of noisy best-reply dynamics is carried out as in Kandori, Mailath and Rob (1993). The way we formalize the stochastic process is slightly different: exploiting the fact that the underlying Markov chain is regular, from the original chain (that will be called \( p(\varepsilon) \) and is defined over \( \Theta \)), we derive a new chain (that we will call \( \tilde{p}(\varepsilon) \)) that is defined over the set of states that are absorbing in the unperturbed process (that is a subset of \( \Theta \)). Doing this has the advantage that, whenever we are able to identify the set of absorbing states of the unperturbed process with the set of Nash-equilibria of the underlying game, we can think of \( \tilde{p}(\varepsilon) \) as ranging over the latter set. This formalizes the intuition that appears in Kandori, Mailath and Rob (1993) and in Canning (1992) for an underlying 2-2 Coordination game. Although we do not pursue this explicitly, we conjecture that, for particular classes of games, the rates at which the transitions in \( \tilde{p}(\varepsilon) \) occur correspond exactly to the radius and coradius identified in Ellison (1995), quoted in Footnote 4. We do not report the equilibrium selection result obtained in that paper, because the result refers to a "Darwinian dynamics" that does not correspond exactly to the best-reply we analyze here. I further conjecture that the extension of that result to best-reply dynamics may be obtained by using the definition of risk-dominance we use in this work, rather than the one introduced by the author. The result of Theorem 43 is only a re-statement of a result of Kandori and Rob (1993), that underlines the fact that the definition of risk-dominance we adopt incorporates many of the requirements needed to in the proof. Example 42 makes, once again in this dissertation, the point that a model of local interaction is to be described by looking at
the original state-space, \( \Theta \), and not at a lumped version of it. Although a \( \tilde{\phi}(\varepsilon) \) can be derived for a locally interactive model in exactly the same fashion as for a population matching model, the set of relevant states, among which to selection is to take place, may radically differ.

The analysis of best-reply dynamics with payoff dependent mistakes relates to Blume (1993) and An and Kiefer (1992). The class of behavioural rules we study is exactly the same. The main difference from Blume (1993) is that, in the model we study, the population is finite. The adjustment process studied in that work relies on a continuous time formulation that can be thought of as a limit, for the time interval becoming infinitesimally small, of the asynchronous dynamics we analyze. The nature of our model is closer to that of An and Kiefer (1992), where players play with all neighbours and take into consideration the average payoff obtained in a round of interaction. Besides the result, I believe the main value added of our model, with respect to other in the same line of research, is that it aims at providing a motivation, in terms of average payoff in the population, to an otherwise exquisitely technical formalization. On one hand, I hope this helps to partially alleviate the sense of frustration that the reader might experience when faced with aseptic techniques. On the other hand, as Example 48 shows, the quantity that appears at the exponent of the formula in Theorem 46 a) is easy to calculate and b) reminds of a correlated equilibrium, which is what we started with in Chapter 1. Having said that, Section 6.3.2 can be skipped without loss of continuity.

6.2 Noisy best-reply Dynamics

We follow the specification of mistakes introduced in Kandori, Mailath and Rob (1993) and in Young (1993). The underlying idea is that players would follow the
myopic best-reply rule defined in Section 3.2 of Chapter 3, but their decision process is perturbed by mistakes. In particular, mistakes affect each player's decision in a way that is independent of any relevant specification of the model: when asked to choose one out of $m$ available actions, a player myopically best-responds with probability $1 - (m - 1) \varepsilon$ and with probability $\varepsilon$ will adopt each (and any) of the remaining actions.

**Definition 38** Given $\Gamma = (\Omega, G, \mu)$, any noisy best-reply dynamics $\varphi(\varepsilon)$ is a process for which the probability with which the transition from state $\theta'$ at time $t - 1$ to state $\theta$ at time $t$ occurs, denoted by $P_\varepsilon(\theta_t | \theta_{t-1})$, is such that for all $\omega$:

$$\Pr(a_{\theta_i}(\omega) = i \in \text{Arg}\max_{i \in \mathcal{A}} \sum_j \pi_{ij} \mu_{\theta_{t-1}}^{\theta_{t-1}}(\omega)) = 1 - (m - 1) \varepsilon < 1$$

$$\Pr(a_{\theta_i}(\omega) = j \notin \text{Arg}\max_{i \in \mathcal{A}} \sum_j \pi_{ij} \mu_{\theta_{t-1}}^{\theta_{t-1}}(\omega)) = \varepsilon > 0 \quad \forall j \neq i$$

where $\varepsilon$ is a constant parameter.

A general specification of the aggregate dynamics under noisy best-reply is hence obtained as a "perturbation" of a myopic best-reply dynamics of which in Chapter 3. As we recall, given a population game $\Gamma = (\Gamma, G, \mu)$ the latter was described by a Markov chain over the state-space $\Theta$ of all possible configurations of play. We denoted its transition matrix by: $\varphi = \begin{bmatrix} I & 0 \\ R & Q \end{bmatrix}$. A noisy best-reply dynamics, $\varphi(\varepsilon)$, is then characterized by a Markov chain, ranging over the same state-space, where each entry of the original transition matrix is perturbed according to the parameter $\varepsilon$. To keep an explicit reference to the underlying noiseless process, we express $\varphi(\varepsilon)$ in the same canonical form:

$$\varphi(\varepsilon) = \begin{bmatrix} T(\varepsilon) & U(\varepsilon) \\ R(\varepsilon) & Q(\varepsilon) \end{bmatrix}$$

Specifically, $T(\varepsilon)$ describes transitions within the set of states that were absorbing under the myopic best-reply dynamics. In Chapter 3, Section 3.3 we denoted this set
as $A^p(\Gamma)$. $U(\varepsilon)$ refers to transitions from states in $A^p(\Gamma)$ to states outside it, i.e. to the set of states that were transient under myopic best-reply dynamics. The analysis of Chapter 3 showed that $A^p(\Gamma) \subset \Theta$ and characterized each element of $A^p(\Gamma)$ in terms of equilibria of the population game $\Gamma$.

By the construction of the transition probabilities, $\lim_{\varepsilon \rightarrow 0} \varphi(\varepsilon) = \begin{bmatrix} I & 0 \\ R & Q \end{bmatrix} = \varphi$, entry by entry. Furthermore, for all $\varepsilon > 0$, $\varphi(\varepsilon)$ is regular, in that all its entries are strictly positive. As such, $\varphi(\varepsilon)$ admits a unique limit distribution (Kemeny and Snell (1976), Theorem 4.1.6). Let this be given by the $\#\Theta$- (row) vector $\alpha(\varepsilon)$. We are interested in characterizing the process when it is observed in any of the states in $A^p(\Gamma)$. Hence, from $\varphi(\varepsilon)$, we derive the matrix describing the transitions (not necessarily in one step) between any two states in $A^p(\Gamma)$ as:

$$\bar{\varphi}(\varepsilon) = T(\varepsilon) + U(\varepsilon)(I - Q(\varepsilon))^{-1}R(\varepsilon)$$

$\bar{\varphi}(\varepsilon)$ formalizes the fact that, starting from any state in $A^p(\Gamma)$, under noisy best-reply, the process can move to another state in $A^p(\Gamma)$ either in one step (and the entry is given by $T(\varepsilon)$), or it can move to a state outside $A^p(\Gamma)$ (with probabilities given by $U(\varepsilon)$). In the latter case the process enters $A^p(\Gamma)$ for the first time with probabilities given by $(I - Q(\varepsilon))^{-1}R(\varepsilon)$ (See Lemma 17 in Chapter 3). It is easily seen that, given that $\varphi(\varepsilon)$ is regular, the same is true for $\bar{\varphi}(\varepsilon)$. Furthermore, its fixed probability vector $\bar{\alpha}(\varepsilon)$, can easily be obtained by re-normalizing the components referring to states in $A^p(\Gamma)$, of the fixed vector of $\varphi(\varepsilon)$ (see Kemeny and Snell (1976), Theorem 6.1.1). We summarize these considerations in the next Remark.

---

1A regular Markov chain, $\varphi$, is such that, there exists a $t$, such that all entries of $\varphi^t$ are strictly positive. Under synchronous updating this is true from time $t = 0$.

2Specifically, $\alpha_\Theta(\varepsilon)$ is a vector defining a probability distribution over the set of states $\Theta$. Any probability vector specifying the initial condition, $P^\Theta(0)$, will converge, as time goes to infinity, to a vector of constants $\alpha_\Theta(\varepsilon)$. We drop the subscript $\Theta$ for notational convenience.
Remark 15 Under noisy best-reply, for any $0 < \varepsilon < \frac{1}{m-1}$, $\bar{\varphi}(\varepsilon)$ is a regular Markov chain over the set of states $A^0(\Gamma)$, with fixed probability vector $\bar{\alpha}(\varepsilon)$.

$\bar{\varphi}(\varepsilon)$ summarizes most of the information contained in $\varphi(\varepsilon)$; its fixed vector, $\bar{\alpha}(\varepsilon)$, provides us with a description of the long run behaviour of the system. We then aim at characterizing the latter. A way of explicitly doing that has been used in the recent literature; we report it in the Lemma that follows:

Lemma 39 Let $\varphi$ be a (finite) regular Markov chain over the set of states $X$, with fixed vector $\alpha$. For each $x \in X$, an $x$-tree, $T_x$, is a subset of directed edges such that, for every vertex $z \in X, z \neq x$, there exists a unique directed path in $T_x$ from $z$ to $x$. Let $T_x$ be the set of all $x$-trees. For each $x$, define the number:

$$P_x = \sum_{T_x \in T_x} \prod_{(x,z) \in T_x} P(x \mid z)$$

where $P(x \mid z)$ is the transition probability from $z$ to $x$. Then:

$$\alpha_x = \frac{P_x}{\sum_{x \in X} P_x}$$

Proof. Freidlin and Wentzell (1984), Lemma 3.3.

Recall that the invariant distribution is parametrized by $\varepsilon$. For the purposes of our analysis, we shall be interested in characterizing its limit for $\varepsilon \to 0^3$. Given each generic entry $P_\varepsilon(\theta' \mid \theta)$, let $r_{\theta'} \geq 0$ be a number such that $0 < \lim_{\varepsilon \to 0} \varepsilon^{-r_{\theta'}}$

---

3Recall that the idea underlying the model is that each player makes a mistake with probability $\varepsilon$. As $\varepsilon \to 0$, players do not make mistakes any further. Given that when they do not make mistakes they behave according to the myopic best reply rule, if the system is stationary, as $\varepsilon \to 0$, they learn to behave optimally. The thought experiment that we, the modellers, do consists of two steps: first we let $t \to \infty$, in order for the (unique) limit distribution to be a good predictor of the probability with which the system is observed in any state; as a comparative statics exercise we then let $\varepsilon \to 0$. In many cases of interest many states disappear from the support of $\alpha(\varepsilon)$; those that remain in the support are said to be selected by the learning process. As it is clear, the whole logic collapses if the order with which limits are taken is reversed.
$P_e(\theta' | \theta) < \infty$. Specifically, $r_{\theta\theta'}$ defines the speed, or the rate, at which each entry of the perturbed process $\varphi(\varepsilon)$ reaches the corresponding entry of the unperturbed process $\varphi$. It is intuitive to think of it in terms of mistakes necessary to move the system from one state to another. By assumption $r_{\theta\theta'} = 0$ for all and only those transitions that occur with positive probability in the unperturbed process. For any $\theta$-tree $T_\theta$, we can then derive the speed at which the system reaches $\theta$, along the tree, simply as the sum of the rates at which transitions on each single branch occur. Let $\sum_{(\theta, \theta') \in T_\theta} r_{\theta\theta'}$ be the latter number. By standard considerations, the limit behaviour of $\sum_{T_\theta} \sum_{(\theta, \theta') \in T_\theta} r_{\theta\theta'}$ will be shaped by the $\min_{T_\theta} \sum_{(\theta, \theta') \in T_\theta} r_{\theta\theta'}$. Following the terminology introduced by Young (1993) we call this quantity the stochastic potential of $\theta$. Then, from the above result, we can infer that the limit of the ratio between any two entries of the invariant distribution, will be determined by the ratio between the corresponding stochastic potentials. This is the logic followed in the next Lemma:

**Lemma 40** Let $\bar{\varphi}(\varepsilon)$ and $\bar{\alpha}(\varepsilon)$ as previously defined. Then:

$$\lim_{\varepsilon \to 0} \bar{\alpha}_\theta(\varepsilon) > 0 \iff \theta \text{ has the minimum stochastic potential.}$$

**Proof.** Young (1993), Lemma 1.

The above result has fruitfully been applied to population models for some classes of games, where the state that shows the minimum stochastic potential is said to be selected under noisy best-reply dynamics\(^4\). A result, now standard, is the following:

---

\(^4\)A different way of characterizing the limit behaviour of the stochastic process has been recently suggested in Ellison (1995). In an entirely analogous framework, the author provides a description of the long run behaviour of the system in terms of measures of the persistence of different absorbing states of the unperturbed process. These are the radius, i.e. the minimum number of mistakes necessary to leave the basin of attraction of a state, and the coradius, i.e. the maximum over all states of the minimum number of mistakes necessary to enter the basin of attraction of a state. The main result of the paper shows that a sufficient condition for a state to be selected, as $\varepsilon \to 0$, is if the radius is greater than the coradius, or a modified version of it. These measures also provide a bound for the speed at which convergence takes place.
Theorem 41 Given $\Gamma = (\Omega, C_2, \mu_U)$, whenever a risk-dominant equilibrium exists, it will be selected under the noisy best-reply dynamics.

Proof. Kandori, Mailath and Rob (1993). An analogous result is obtained for "adaptive play (where players best-respond to a sample of observations of past play) with mistakes" by Young (1993).

It has been argued (Ellison (1993), (1995)) that Theorem 41 is robust to the specification of the matching technology. In the models analyzed there, players are located on a circle or on a square lattice respectively, and interact only with their "nearest" neighbors. In what follows we apply the same line of argument to an example of a simple model of local matching, for which instead the result does not hold.

Example 42 Consider $\Gamma = (\Omega, C_2, \mu)$, where $\#\Omega = 4$, $C_2 = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & 1 \end{bmatrix}$ and $\mu$ is specified as in the following graph, where vertices are the players, identified by $a, b, c, d$, edges are elements in the support of $\mu$, and the number above each edge is the probability with which each matching occurs:

```
 a 1/3 b
1/6 1/6
 c 1/3 d
```

We label actions available to players as 1 and 0. It is not difficult to show that the following configurations of play are absorbing states under myopic best-reply dynam-

---

5In the aforementioned paper, equilibrium selection results are obtained for all classes of 2-2 symmetric normal form games (and not limited to coordination games). Rhode and Stegeman (1996) recently pointed out that for games other than coordination games, the results do not necessarily hold for finite population games.
For (C) and (D), note that, for example player a is likely to be matched with player b, who plays exactly the same action as he does, with conditional probability equal to $\frac{2}{3}$, and with player c with probability $\frac{1}{3}$. Given that $\frac{2}{3} > \frac{1}{3}$, its expected payoff is maximized by choosing action 1 in configuration (C). In (C), by the same token player c is choosing 0 as his strict best-reply (because $1\frac{2}{3} > \frac{1}{2}\frac{1}{3}$). We now aim at characterizing the limit distribution under noisy best-reply. The state space of the Markov process is given by the $4^2$ possible configurations of play. After tedious, but straightforward computations it can be shown that the rates at which transitions occur are the following:

Each $r$ describes the minimum number of mistakes necessary to move between states (given that transitions allowed in the unperturbed process have a cost of 0, the latter is simply the number of mistakes necessary to get from one absorbing state into the
basin of attraction of the other). It is now not difficult to see that, for absorbing state (C) for example, the minimum stochastic potential is achieved on a tree of the kind:

The same reasoning extends to every absorbing state in exactly the same fashion. As a result, the limit distribution will place equal probability on all absorbing states. Hence, though equilibrium (1, 1) is risk-dominant, it will not be selected by the dynamics. Note that the result is not due to the fact that matching is not uniform; entirely analogous conclusions hold for a model that differs from the above one in that 6 players are matched, according to \( \mu_L \), in the following way:

Though simple, the above example underlines potential problems that one may encounter in applying the logic underlying Lemma 40 to settings that are more com-
plex than that of Theorem 41. First, the identification of the absorbing states of the unperturbed process: the latter is trivial if the underlying game is a 2-by-2 coordination game, but not necessarily so otherwise. In some cases (a leading example being the topology of the circle with nearest neighbor interaction and an underlying 2-by-2 coordination game, as in Ellison (1993)), it is easy to identify the set of absorbing states of the unperturbed dynamics with the set of equilibria of the underlying game. As pointed out in Chapters 1 and 3, this is however not true in general.

Second, even if we are able to identify absorbing states, unless the underlying game is 2-by-2 and matching is uniform within the whole population, basins of attractions of different absorbing states may overlap. This complicates the study because when evaluating the minimum speed at which a transition between two different absorbing states may occur, we need to consider all possible paths that achieve a state by passing through potentially all different absorbing states. For instance, in the above Example the number of mistakes necessary to move from (A) to (B) by a direct jump is exactly the same as the number of mistakes necessary to move first from (A) to (C) and then from (C) to (B). As Kandori and Rob (1993) point out, even in simple 3-by-3 coordination games it may happen that direct jumps are even more costly, in terms of mistakes, than indirect transitions. Though in principle all these computations may be done, it is however difficult to identify equilibria that are robust to this kind of perturbation simply by looking at basic properties of the underlying game. A different way of stating the same remark, is that, as noted in Chapter 3, basins of attractions under myopic best-reply may, in general, overlap.

Lastly, the generalization of the result that risk-dominance constitutes a sufficient condition for equilibrium selection in generic $m$-$m$ coordination games under uniform population matching, requires further assumptions on the geometry of the best-reply regions. Along these lines, Kandori and Rob (1993) provide conditions under which an equilibrium that pairwise risk-dominates any other, will be selected by the noisy best-reply dynamics. The conditions require the game to satisfy specific properties
referred to as bandwagon properties. The result that follows builds upon that line of argument.

\[ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \]

**Theorem 43** Given \( \Gamma = (\Omega, C_3, \mu_U) \), assume \( \pi_{ij} = \pi_{ij} \; \forall l \neq i \neq j \). Then, whenever a risk-dominant equilibrium exists, it will be selected under the noisy best-reply dynamics.

**Proof.** First notice that if \( \pi_{ij} = \pi_{ij} \; \forall l \neq i \neq j \), then, by Remark 4 in Chapter 1, absorbing states of the myopic best-reply dynamics are all and only those states where all players adopt exactly the same action, and by Theorem 14 in Chapter 3, the unperturbed process is proved to converge. The proof is simply an application of the result of which in Kandori and Rob (1993). It amounts to show that if \( C_m \) is as above, then the monotone share property (MSP) and the total bandwagon property (TBP) are satisfied. TBP requires that for any probability distribution over actions \( p \), the best-reply to \( p \) be in the support of \( p \). TBP implies that, for any subset of actions \( A' \subseteq A \), there exists a unique mixed strategy equilibrium with support \( A' \). If TBP holds, then MSP requires the following: for any two subsets of actions \( A' \) and \( A'' \) such that \( A' \subset A'' \), let \( p' \) and \( p'' \) be two probability distributions with support \( A' \) and \( A'' \) respectively. Then MSP holds if \( p'_{ij} > p''_{ij} \) for all \( j \in A' \). Theorem 1 in Kandori and Rob (1993) shows that if a coordination game satisfies TBP and MSP, then the minimum number of mistakes required to jump from equilibrium \( \theta_i \) to equilibrium \( \theta_{j \neq i} \) is achieved by a "direct jump" from \( \theta_i \) into the basin of attraction of \( \theta_j \).

To show that the above properties hold in our contest we shall use the geometry of a simplex. Consider \( C_3 \) and assume equilibrium \( i \) is risk-dominant as in Definition
20 in Chapter 3. Then the best-reply regions must look as follows:

![Diagram](image)

To see this note that if \( \pi_{ij} = \pi_{lj} \quad \forall l \neq i \neq j \), any pair of actions performs equally well towards a probability distribution that puts mass one on the third action. This implies that, for example, the locus of probability distributions towards which actions \( i \) and \( j \) yield the same expected payoff must include point \( l \). Furthermore, given that the property in Remark 4 of Chapter 1, is a direct implication, the slope of the borders of the best-reply regions must be as in the above picture: any slope that does not lie in the angle exiting from \( a \) would in fact contradict the assumption. This suffices to show that TBP and MSP are satisfied. Given that equilibrium \( i \) is, by assumption, risk-dominant, \( ia \) and \( ib \) must in turn be greater than \( \frac{1}{2} \). But this coincides with the definition of risk-dominance of which in Kandori and Rob (1993). The conclusion of the proof is then a straightforward application of Theorem 2 in that paper.

A peculiar feature of the results obtained by using the above line of argument is that in most cases, conclusions rely on the specific features of the underlying game. Moreover, a substantial limit of the approach has been underlined in Bergin and Lipman (1996), where the authors show that if mistakes are state-dependent (in that they occur at different rates in different states), then the above line of argument fails to produce unique predictions. In particular, any of the limit distributions of the unperturbed process, may be selected by an appropriate choice of mutation rates.
6.3 Payoff-dependent Mistakes.

As an introduction to the model we are going to analyze in this section consider the following. Assume the underlying game is a 2-by-2 coordination game and, given a state as initial condition, the best-reply is the same for all players in the population. If each player adopts the noisy best-reply, as specified in the previous section, then he will choose the myopic best-reply, say action 1, with probability $1 - \varepsilon$, and action 0 with probability $\varepsilon$. Given that the same reasoning will apply to all subsequent periods, is behaviour is fully described by a regular Markov chain over his action space $A = \{1, 0\}$. The aggregate population can hence be thought of as a finite collection of $\#\Omega$ mutually independent Markov chains. The analysis of this system is entirely elementary because of the independence assumption and the simple nature of the individual chains. If we view the system as a Markov process over $A^\Omega$, for any fixed initial condition $\theta_0$, the distribution of $\theta_t$ at different times are product measures which are mutually singular with respect to each other and with respect to the unique invariant measure of the process. Hence the unique limit distribution will be exactly the product of the invariant distribution of the individual two state Markov chain. This is due to the fact that, once an initial condition is specified, choices that players make, do not depend on the interaction: the initial condition defines expected payoffs and these (or the ranking between them) are exactly the same for all players. In this section we aim at analyzing a model where, instead, interaction plays a role in determining the evolution of the aggregate system. The way we do that relies on two key assumptions: a) the probability with which actions are chosen does explicitly depend on expected payoffs, b) expected payoffs differ among players. Due to b), although the process may maintain Markovian properties, each single player behaviour is no longer Markovian: the probabilities with which player $\omega$ chooses each available action depend in fact on the actions chosen by his potential opponents. As for a), we shall assume that the functional form that defines the relation between action choices and expected payoffs is time independent and
exactly the same for all players.

6.3.1 The Model

In the line of the payoff-dependent stochastic updating rule introduced in Chapter 2, we specify players behaviour as follows:

$$\Pr(a_t(\omega) = i \in A) = \frac{\exp[\sigma \pi^t_u(\omega)]}{\sum_{i \in A} \exp[\sigma \pi^t_u(\omega)]}$$

where, we recall, $\pi^t_u(\omega)$ is the payoff that player $\omega$ expects to get from the round of interaction at time $t$, if he chooses action $i$. $\sigma$ is a non negative parameter that we use to model mistakes in the following way. If $\sigma = 0$, then player $\omega$ is equally likely to choose any of the available actions, independently of any payoff consideration. For any $\sigma > 0$, player $\omega$ still adopts a completely mixed strategy, but the probability with which each action is chosen is strictly increasing in the expected payoff achievable with that action. Hence, he is more likely to choose a best-reply to a given configuration of play, then any other action. Moreover, if he does not best-respond, i.e. if he makes a mistake, he is less likely to make very costly mistakes, that is mistakes that entail a high loss in his expected payoffs. As $\sigma$ increases, player $\omega$ becomes more and more likely to adopt the best-reply (and less and less likely to make mistakes). In the limit, for $\sigma \to \infty$, player $\omega$ will only adopt a best-reply, i.e. he will not make any mistake. Given that the above formalization is continuous in $\sigma$, one feature of the model is that, by an appropriate choice of $\sigma$, one can get an arbitrarily close approximation of best-reply behaviour.

As previously stated, we want to allow for expected payoffs to differ across players; this is for example the case if we consider a model of local uniform matching, as in Definition 10 in Chapter 3, where, at each time $t$, the probability with which player

---

*While under noisy best-reply a player wanting to go to the 20th floor by lift, is equally likely to press button 19 and button 1, under payoff dependent mistakes, the probability with which he presses buttons is decreasing in the number of levels below floor 20.*
\( \omega \) is matched with player \( s \) choosing action \( l \), is simply the number of \( l \)-players, at time \( t - 1 \), in the set of \( \omega \)'s potential opponents. If we let \( |\Omega_\omega| = n \) for all \( \omega \), and \( |\Omega|^l \) denote the number of player \( \omega \)'s potential opponents observed choosing action \( l \) by player \( \omega \), the above rule becomes:

\[
\Pr(a_\omega(s) = i | s \in \Omega_\omega) = \frac{\exp[\sigma \sum_{i \in A} \frac{|\Omega|^l_{i-1}}{n}]}{\sum_{i \in A} \exp[\sigma \sum_{i \in A} \frac{|\Omega|^l_{i-1}}{n}]} \tag{6.1}
\]

To lighten notation, we refer to (6.1) as \( \pi_\omega(|\Omega|^l_{i-1}) \) (i.e. the probability law with which player \( \omega \) adopts each action \( i \), given the actions chosen in his neighbourhood) and we drop the time subscript whenever not explicitly needed. We shall specify the dynamics in terms of random asynchronous updating: at each time \( t \) only one player, chosen at random in the population, is allowed to revise his action. He will do so, by adopting each action \( i \) with probability given by (6.1).

**Definition 44** Given \( \Gamma = (\Omega, G, \mu_L) \), any asynchronous best-reply dynamics with payoff dependent mistakes, \( \varphi(\sigma) \), is a process for which, if (and only if) state \( \theta \) differs from state \( \theta' \) only by the action of player \( \omega \), who adopts action \( i \) in state \( \theta \), then the probability with which the transition from \( \theta' \) at time \( t - 1 \), to \( \theta \) at time \( t \) occurs is given by:

\[
P_\sigma(\theta_t | \theta'_{t-1}) = \frac{1}{|\Omega|^l} \pi_\omega(i_t | \Omega^l_{i-1}).
\]

In line with the analysis of the previous part of this Chapter, we are interested in describing the long run behaviour of the process and in explicitly characterizing it. Many of the considerations we stated in Section 6.2 apply to this process as well. In particular, this process too is described by a regular Markov chain, as stated next.

**Remark 16** Under myopic best-reply with payoff dependent mistakes, for any \( 0 < \sigma < \infty \), \( \varphi(\sigma) \) is a regular Markov chain over the set of states \( \Theta \). We denote its fixed probability vector by \( P_\sigma(\theta) \) for \( \theta \in \Theta \).
Proof. From Definition 44, for each finite $\sigma$, transition probabilities are time independent and strictly positive. This together with the fact that $\Theta$ is a finite set proves the claim.

In what follows we shall explicitly characterize the limit distribution of the process, by using the conditional probability of which in (6.1). For simplicity, while doing that, the parameter $\sigma$ is taken to be equal to one and will explicitly re-appear as $\sigma$ only at the end.

6.3.2 A "local" characterization of the process

We regard each player's action as a random variable, $a(\omega)$, with support $A$ and probability distribution given by (6.1). The family $\{a(\omega), \omega \in \Omega\}$ will be assimilated with an element in $\Theta = A^0$. As previously done, elements of $\Theta$ will be called configurations of play. We shall think of any probability distribution over the state space $\Theta$ as a random field\footnote{In general, if $\Psi$ is a denumerable set, $\mathcal{F}$ is the collection of all subsets of $\Psi$, and $P$ is a probability measure on $\mathcal{F}$, the triple $(\Psi, \mathcal{F}, P)$ is called a random field.}, referred to as $P$ in what follows. $P$ is thus a probability measure over the set of all mappings that assign an action to each player.

We aim at describing the probability distribution of $P$ in terms of our knowledge of the collection of all $a(\omega)$. We first specify $P$'s domain. We take $\Omega$ to be a finite graph, for which $\omega$ and $s$ denote two generic sites. If $\omega$ and $s$ are connected by an edge, we shall call them neighbours. The typical example we shall work with is the boundaryless structure of the $\Omega$-Torus that we described in Chapter 4. In this case $\Omega$ is connected, in the sense that, for $\omega$ and $s$ in $\Omega$, there exists a finite path of edges joining $\omega$ and $s$ and $\|\omega_\omega = \|\omega_s = n$ for all $\omega$ and $s$ (i.e. the graph is $n$-regular). We find it useful to identify neighbourhoods in terms of the Euclidean distance, that we denote by $|\cdot - \cdot|$. On the $\Omega$-Torus, we shall then specify $\Omega_\omega = \{s : |\omega - s| \leq \sqrt{2}\}$. Most of the analysis does, however, carry through for any denumerable $\Omega$ equipped with a distance, for which a boundary value function is given. We then require $P$ to
be consistent with the system of conditional probabilities \( \{p_\omega(i \mid \Omega_\omega), \omega \in \Omega\} \). By that we mean that the latter can be obtained from \( P \), as \( P(a(\omega) = i \mid a(\cdot) = f(\cdot) \text{ on } \Omega - \{\omega\}) \), that is by conditioning over the actions adopted by all other players in the population. Whenever we are able to do so, then we characterize the process as in the following Remark:

**Remark 17** Let \( \Omega \) be an \( \Omega \)-Torus, \( \Theta = A^\Omega \), \( \mathcal{F} \) be the collection of all subsets of \( \Theta \) and \( P \) be a probability measure on \( \Theta \), such that for all players \( \omega \) and for any action \( i \), \( \Pr(a(\omega) = i) = p_\omega(i \mid \Omega_\omega) \) (where \( p_\omega(i \mid \Omega_\omega) \) is given by (6.1)). Then the triple \( (\Theta, \mathcal{F}, P) \) is a Markov random field on the \( \Omega \)-Torus.

**Proof.** A Markov random field is a random field for which the following properties hold: a) \( P(\theta) > 0 \) for all \( \theta \in \Theta \) (positivity), b) conditional probabilities of the form \( P(a(\omega) \mid a(\cdot) = f(\cdot) \text{ on } \Omega - \{\omega\}) \) depend only on the values of \( f \) in a neighbourhood of \( \omega \) (i.e. the points within a given distance from \( \omega \)), c) the above conditional probabilities are translation invariant (i.e. \( P(a(\omega) \mid a(\cdot) = f(\cdot) \text{ on } \Omega - \{\omega\}) = P(a(s) \mid a(\cdot) = g(\cdot) \text{ on } \Omega - \{s\}) \) whenever \( f(\omega + z) = g(s + z) \) for all \( |z| \leq r \). These properties are satisfied by Definition 6.1: given that each action can be adopted with strictly positive probability, the field is positive; conditional probabilities depend only on a the configuration of play within a player's neighbourhood; lastly, they are translation invariant in that expected payoffs are linear in the number of players adopting each action, hence players facing the same distribution of actions in their respective neighbourhood, will adopt exactly the same mixed strategy. We ignore a boundary condition in that \( \Omega \) is assumed to be a Torus. 

A useful way of explicitly deriving \( P \)'s probability distribution, borrows techniques extensively used in physics, that we review in the Subsection that follows.
Energy Functions and Gibbs Measures.

The underlying idea is to specify a random field of the kind we introduced above, in terms of an energy function (alternatively referred to as a potential or an Hamiltonian) that is a particular function of the configuration of system. The existence of an energy function guarantees that the system of conditional probabilities satisfies the consistency condition. A random field defined in terms of a potential is known as a Gibbs measure; the problem of identifying conditions under which Gibbs measures exist and they are unique is known as the D.L.R. problem.

We focus on a class of random fields specified in terms of an interaction potential $I$. Given any subset of players, $V \subseteq \Omega$, let $\prod_{\omega \in V} a(\omega) = \Theta(V)$ be the configuration of play in the subset $V$. $I$ is then a mapping that associates a real number to each $\Theta(V)$: $I_V : \Theta(V) \rightarrow \mathbb{R}$. We construct an interaction potential by taking a real function $J(a(\omega), a(s)) : A \times A \rightarrow \mathbb{R}$, that associates a number to each combination of actions chosen by the pair of players $\{\omega, s\}$. We then let $I_{\{\omega, s\}}(a) = J(a(\omega), a(s))$. The total potential, or energy, on a finite subset $W$ of interactions will then be:

$$H_W(a) = \sum_{W \cap V \neq \emptyset} I_V(a) = \sum_{\{\omega,s\} \in W} J(a(\omega), a(s))$$

The contribution of each player $\omega$, to the energy of the subset $V$, is defined as $U_\omega(a) = \sum_{\omega \in V} \frac{1}{I_V} I_V(a)$ whenever the sum is finite. Finally, we let $H(a)$ (the Hamiltonian) denote the energy of the set of all interactions: $H(a) = \sum_V I_V(a) = \sum_\omega U_\omega(a)$.

Given this specification, we shall say that a random field $(\Theta, F, P)$ on the bound-

---

8The probabilistic formulation we adopt in this Section relies on Prum and Fort (1991). I found the latter a good general formalization of a topic that has been studied by mathematicians, mostly probabilists, and to an even greater extent by physicists over the last decade. The bi-dimensional model corresponds to the so-called Ising model of interacting particles, analyzed in detail in Liggett (1985). As a non-technical reference, I found Amit (1989) an excellent descriptive introduction to the field.

9We can also account for what may be called an imposed field, i.e. a potential that is independent of the interaction: $I_{\{\omega\}}(a) = h(a(\omega))$, where $h(a(\omega)) : A \rightarrow \mathbb{R}$. In this case the total energy of a
aryless domain $\Omega$ is a Gibbs random field with interaction potential $I$, if $P$ is defined by the Gibbs formula:

$$P(\theta) = \frac{1}{Z} \exp[-H(a)]$$  \hspace{1cm} (6.2)

where $Z$ is a normalizing constant, for which $\sum_{\theta} P(\theta) = 1$. Note that the dependence on $\theta$ of the r.h.s. stems from the fact that $a : \Omega \rightarrow A^\Omega = \Theta$, in that what is observed of $\Omega$ (the set of all players) is actually a configuration of play (an element of $\Theta$). Physicists attribute to this expression the following meaning: if $H(a)$ was known for every state of the system, then the properties of the system, for a given parameter $\sigma$ (set equal to one for simplicity in the above expression), could be computed as if we had an ensemble of identical systems and the probability for finding one of them in any of the possible states is proportional to $H(a)$.

We shall refer to the set of Gibbs measures (or the Gibbs ensemble) associated with the interaction potential $I$ as $\{G_I\}$. We shall use the above formalization to model the dynamics of the system. In particular, our interest in these techniques stems from the fact that we shall characterize the limit distribution of the process under myopic best-reply with payoff dependent mistakes as a Gibbs measure. In what follows, we first focus on the class of interaction potentials (and corresponding Gibbs measures) appropriate to our model; we then show that a random field constructed in this way fulfills the Markovian requirements that we defined in the previous Section. Finally, we specify sufficient conditions to guarantee ergodicity of the process. The idea is to relate the analysis to that of Chapter 3, where, we recall, we studied the noiseless version of the models we analyze in this Chapter.

Subset of interactions will be:

$$H_W(a) \equiv \sum_{W \cap V \neq \emptyset} I_V(a) = \sum_{\{\omega, s\} \in W} J(a(\omega), a(s)) + \sum_{\omega \in W} h(a(\omega)).$$
Given the nature of the model we analyze, we restrict the attention to a specific class of interactions. In particular, we look at pairwise interactions, for which \( \|V \| > 2 \Rightarrow \mathcal{I}_V(\cdot) = 0 \), of finite range, i.e., interactions for which there exists an \( r \) s.t. \( \text{diam}(V) > r \Rightarrow \mathcal{I}_V(\cdot) = 0 \). The latter assumption suffices to ensure that \( \mathcal{U}_\omega(a) \) is finite. It is immediate to notice that the specification of the interaction pattern we adopted throughout this work incorporates the first requirement, in that \( \mu \) is defined over couples of players. Furthermore, if matching is local, also the second requirement is fulfilled, in that for a given \( \omega \), \( \mu(s \mid \omega) \) has support \( \Omega_\omega \subset \Omega \). We shall explicitly use \( \mu \) to formalize the class of interactions we analyze, by letting \( \mathcal{I}_{(\omega,s)}(a) = J(a(\omega), a(s))\mu(\{\omega, s\}) \). Lastly, note that we are studying a model of uniform local matching, where, by definition, \( \mu(\{\omega, s\}) \) is constant. Hence the interaction potential of a couple of players depends only on the action adopted by the two players (and not on their specific identity). This feature ensures that the potential is translation invariant. We denote any translation invariant, pairwise interaction potential of finite range as \( \mathcal{I}^L \). First we show that any Gibbs random field constructed in this way, is a random Markov field as defined in Remark 17.

**Theorem 45** Every Gibbs random field with the interaction potential \( \mathcal{I}^L \) is a Markov random field.

**Proof.** We start by taking a Gibbs random field and proceed to verify conditions (b) and (c) stated in Remark 17 (condition (a) is obvious). By standard arguments, conditional probabilities are given by:

\[
P(a(\omega) = i \mid a(\cdot) = f(\cdot) \text{ on } \Omega - \{\omega\}) = \frac{P(a(\omega) = i \text{ and } a(\cdot) = f(\cdot) \text{ on } \Omega - \{\omega\})}{\sum_{i \in A}[P(a(\omega) = i \text{ and } a(\cdot) = f(\cdot) \text{ on } \Omega - \{\omega\})]}
\]

For a Gibbs measure, from 6.2 we know that the numerator is:

\[
\frac{1}{Z} \exp[-H(a \mid a(\omega) = i)] = \]

128
This is also the first term of the sum at the denominator. Any other term of the sum can be obtained in an analogous fashion, simply by substituting \( a_j \neq i \) (instead of \( i \)) in the above formula. Hence conditional probabilities can be written as:

\[
P(a(\omega) = i | a(\cdot) = f(\cdot) \text{ on } \Omega - \{\omega\}) = [1 + \sum_{j \neq i} \exp\left(\sum_{\{s, \omega\}} J(f(s), i)\mu(\{\omega, s\}) \right) - \sum_{\{s, \omega\}} J(f(s), j)\mu(\{\omega, s\})]^{-1}
\]

It is now easy to notice that, given that the interaction potential has finite range, conditional probabilities as above depend only on a finite set of neighbors of \( \omega \), that is the set of \( s \) within distance \( r \) from \( \omega \) (the set \( \Omega_\omega \)). Hence condition (b) is satisfied. (c) requires conditional probabilities to be homogeneous, but this is indeed the case, because the term at the denominator is linear in the configuration of play in \( \omega \)'s neighborhood.

### 6.3.3 Limit behaviour of the process.

As previously stated, we would like to construct a potential on the basis of the payoff matrix of the underlying game. The most intuitive way of constructing a potential is to relate the energy of the interaction between two players to the sum of the payoff achievable by each player, i.e. \( [\pi(a(\omega), a(s)) + \pi(a(s), a(\omega))] \). For convenience, we shall set

\[
J(a(\omega), a(s)) = -\frac{1}{4}[\pi(a(\omega), a(s)) + \pi(a(s), a(\omega))]
\]
(the minus sign is useful to relate the analysis to that of previous Chapters). As a result, the potential associated to the couple of players \(\{\omega, s\}\) is given by:

\[
\mathcal{I}_{(\omega, s)}(a) = 2\mathcal{U}_{(\omega)}(a) = -\frac{1}{4}[\pi(a(\omega), a(s)) + \pi(a(s), a(\omega))]\mu(\{\omega, s\})
\]

and (the negative of the) corresponding Hamiltonian, by:

\[
-H(a) = \sum_{\{\omega, s\}} \frac{1}{4}[\pi(a(\omega), a(s)) + \pi(a(s), a(\omega))]\mu(\{\omega, s\}).
\]

Any associated Gibbs measure, i.e. any element of the set \(\{\mathcal{G}_I\}\) will then be defined by \(P(\theta) = \frac{1}{Z} \exp[-H(a)]\). By re-introducing the parameter \(\sigma\), in \(J(a(\omega), a(s))\) and consequently in \(H(a)\), we obtain: \(P_{\sigma}(\theta) = \frac{1}{Z} \exp[-\sigma H(a)]\), that is an element of \(\{G_\sigma^2\}\).

We shall focus on a class of symmetric normal form games, for which we can characterize the Hamiltonian in terms of (the less aseptic notion of) the average expected payoff in the population. The class of games for which this can easily be done is the class of symmetric normal form games that are balanced, as in Definition 12 in Chapter 3. For this class of games, a Gibbs measure constructed as described above constitutes the unique limit distribution for the aggregate population dynamics under best-reply with payoff dependent mistakes, as proved in the result that follows.

**Theorem 46** Given \(\Gamma = (\Omega, G, \mu_T)\), if \(G\) is balanced, then the following Gibbs measure constitutes the unique limit distribution under myopic best-reply dynamics with payoff dependent mistakes \(\phi(\sigma)\), for any \(0 < \sigma < \infty\):

\[
P_{\sigma}(\theta) = \frac{1}{Z} \exp[\sigma \frac{1}{2} \Pi(\theta)].
\]

**Proof.** Recall that \(\Pi(\theta)\) is the average expected payoff in the population, and is given by:

\[
\sum_{\{\omega, s\}} \pi(a(\omega), a(s))\mu(\{\omega, s\}) = \sum_{\omega} \{ \sum_{s \in \Omega_\omega} \pi(a(\omega), a(s))\mu(s | \omega)\} \mu(\{\omega\}).
\]

130
Clearly, if the game is balanced, then \( \pi(a(\omega), a(s)) = \pi(a(s), a(\omega)) \) for all \((a(\omega), a(s))\). Hence, the Gibbs measure obtained by constructing the interaction potential \( I \) as described above, is exactly \( P_{\sigma}(\theta) = \frac{1}{Z} \exp[\sigma \frac{1}{2} \Pi(\theta)] \). Following Remark 45, this is also a random Markov field over the domain \( \Omega \). We now need to show that such a measure is invariant over time, for the process \( \rho(\sigma) \) for any finite \( \sigma \), set, for simplicity, equal to 1. This is the case if \( P(\theta_t | \theta_{t-1}) P(\theta') = P(\theta'_t | \theta_{t-1}) P(\theta) \). Now suppose that \( \theta \) and \( \theta' \) differ only by the action of player \( \omega \), who adopts action \( j \) in state \( \theta \) and action \( i \neq j \) in state \( \theta' \). Then the ratio between \( P(\theta') \) and \( P(\theta) \) is given by:

\[
\frac{P(\theta')}{P(\theta)} = \exp\left[\frac{1}{2} (\Pi(\theta') - \Pi(\theta))\right] = \exp\left[\frac{1}{2} \sum_{s \in \Omega_\omega} \{\pi(a_{\sigma'}(\omega), a_\theta(s)) - \pi(a_\sigma(\omega), a_\theta(s)) + \pi(a_\sigma(s), a_{\sigma'}(\omega)) - \pi(a_\sigma(s), a_\sigma(\omega))\}\right] = \exp\left[\frac{1}{n} \sum_{s \in \Omega_\omega} \{\pi(a_{\sigma'}(\omega), a_\theta(s)) - \pi(a_\sigma(\omega), a_\theta(s))\}\right] = \exp\left[\frac{1}{n} \sum_{I \in A} (\pi_I - \pi_{ji}) \frac{\#\Omega^I}{n}\right]
\]

The second equality stems from the definition (recall Lemma 11 in Chapter 3) and from the assumption of uniform local matching, implying that \( \mu(s | \omega) = \frac{1}{n} \) \((n = \#\Omega_\omega)\). The third is due to the assumption that \( \pi(a(\omega), a(s)) = \pi(a(s), a(\omega)) \) for all \((a(\omega), a(s))\). The last equality is simply a re-labelling with \( a_{\sigma'}(\omega) = i \) and \( a_\theta(\omega) = j \). \( \#\Omega^I \) denotes the number of players, in \( \omega \)'s neighbourhood, adopting action \( I \). Clearly, the last expression is the relative advantage, for player \( \omega \), of action \( i \) with respect to action \( j \). By definition, that is also the ratio between the one step

\[10\] This condition, known as detailed balance condition, can easily be derived from the general formulation of the time evolution of the process that we gave in the last Section of Chapter 2.
transition probabilities \( P(\theta'_t \mid \theta_{t-1}) \) and \( P(\theta_t \mid \theta'_{t-1}) \):

\[
P(\theta' \mid \theta) \cdot P(\theta \mid \theta') = \frac{1}{\|\Omega\}} \frac{\exp[\sum_{i \in A} \pi_i \frac{\Omega^i_j}{n}]}{\frac{\sum_{i \in A} \exp[\sum_{i \in A} \pi_i \frac{\Omega^i_j}{n}]\|\Omega\}} = \exp[\sum_{i \in A} (\pi_i - \pi_j) \frac{\|\Omega^j_i\|}{n}]
\]

Hence the Gibbs measure constitutes a time invariant distribution for the process \( \varphi(\sigma) \) (for any finite \( \sigma \)). This, together with Remark 16, proves the assertion.

The above result is to be interpreted in the following way. Recall that in Section 3.3 in Chapter 3 we specified conditions under which the average expected payoff in the population was strictly increasing, and as such could be taken to represent a Lyapunov function (or an Energy function) for the system under myopic best-reply dynamics. A sufficient condition for this to be the case is for the underlying game to be balanced. In this case at each step of the noiseless process the average expected payoff is increased. As previously pointed out, the (noiseless) myopic best reply dynamics can be obtained for the model we analyze in this Section with \( \sigma = \infty \). Mistakes occur for any finite value of \( \sigma \). In this case the system can loose energy, at random, in amounts that are proportional to the value of \( \sigma \). Hence some of the local maxima of the energy function, i.e. the average expected payoff, can be destabilized. If actions are chosen in an entirely random fashion, that is for \( \sigma = 0 \), then the system wanders randomly among all its possible states. As a result, for any finite value of the parameter \( \sigma \), we cannot identify states as attractors for the dynamics. However, we can identify attractor probability distributions over states. We interpret the above result as characterizing the class of limit distributions under best-reply with payoff dependent mistakes.
Although the interpretation is less immediate, the result can be generalized. In fact, the sufficient condition of Theorem 46 was designed to interpret the probabilistic formulation we deal with, in terms of the average expected payoff of the game. However, consider a 2-2 underlying game and decompose the payoff matrix as 

\[
\pi = [\hat{\pi}_{ij}] + [\hat{\pi}_i],
\]
given by:

\[
\begin{bmatrix}
\pi_{00} & \pi_{01} \\
\pi_{10} & \pi_{11}
\end{bmatrix} = \begin{bmatrix}
\pi_{00} - \pi_{01} & 0 \\
0 & \pi_{11} - \pi_{10}
\end{bmatrix} + \begin{bmatrix}
\pi_{01} & \pi_{01} \\
\pi_{10} & \pi_{10}
\end{bmatrix}
\]

The first part, \([\hat{\pi}_{ij}]\), clearly reproduces a balanced game. We take the second, \([\hat{\pi}_i]\), to quantify a potential, independent of the interaction, that depends only on the action adopted. We then set:

\[
-H(a) = -\sum_V I_V(a) = \frac{1}{2} \left\{ \sum_{\{\omega,s\}} \hat{\pi}(a(\omega), a(s))\mu(\{\omega,s\}) + \sum_{\omega} \hat{\pi}(a(\omega))\mu(\{\omega\}) \right\}
\]

The first term of the sum is an element of \([\hat{\pi}_{ij}]\) and is to be interpreted in exactly the same fashion as before; the second is an element of \([\hat{\pi}_i]\) that accounts for the security level associated to each action. If the game is balanced, this latter term may be zero. While such a decomposition can always be specified for an underlying 2-2 game, this is not necessarily the case for a generic \(m\text{-}m\) game. Whenever we are able to decompose each entry of the payoff matrix of a generic underlying game in the sum of a quantity that symmetrically depends on the interaction \((n_{ij})\) and a quantity that depends only on each single action adopted \((\hat{\pi}_i)\), then the following result generalizes Theorem 46.

**Theorem 47** Given \(\Gamma = (\Omega, G, \mu_T)\), if \(G\) can be written as \(\pi = [\hat{\pi}_{ij}] + [\hat{\pi}_i]\), then the following Gibbs measure constitutes the unique limit distribution under myopic
best-reply dynamics with payoff dependent mistakes $\varphi(\sigma)$, for any $0 < \sigma < \infty$:

$$P_\sigma(\theta) = \frac{1}{Z} \exp[\sigma \sum_{\omega,s} (\hat{\pi}(a_\theta(\omega), a_\theta(s))\mu(\{\omega, s\}) + \sum_\omega \hat{\pi}(a_\theta(\omega))\mu(\{\omega\})].$$

where $\hat{\pi}(a(\omega), a(s)) \in [\hat{\pi}_{ij}]$ and $\hat{\pi}(a(\omega)) \in [\hat{\pi}_{i}]$.

**Proof.** Time invariance of the above Gibbs measure is proved is exactly the same fashion as in Theorem 46. Indeed, suppose that $\theta$ and $\theta'$ differ only by the action of player $\omega$, who adopts action $j$ in state $\theta$ and action $i \neq j$ in state $\theta'$. Then the following holds:

$$P(\theta') = \frac{1}{P(\theta)} \exp[\sum_{\omega} (\hat{\pi}(a_\theta(\omega), a_\theta(s)) - \hat{\pi}(a_\theta(\omega), a_\theta(s))) + \hat{\pi}(a_\theta(\omega)) - \hat{\pi}(a_\theta(\omega))] =$$

$$= \exp[\sum_{i\in A} (\hat{\pi}_i - \hat{\pi}_j)] \frac{\omega^i}{n + \hat{\pi}_i + \hat{\pi}_j} =$$

$$= \exp \left[ \frac{1}{\omega^i} \exp[\sum_{i\in A} \hat{\pi}_i] \right] = \frac{P(\theta' | \theta)}{P(\theta | \theta')}$$

Uniqueness is guaranteed by Remark 16. 

We conclude this Section with some Examples.

**Example 48** Given $\Gamma = (\Omega, G_2, \mu_\Gamma)$, let $G$ be balanced. The limit behaviour of the corresponding $\varphi(\sigma)$ is characterized by:

$$P_\sigma(\theta) = \frac{1}{Z} \exp[\sigma \sum_{\omega,s} \pi(a_\theta(\omega), a_\theta(s))\mu(\{\omega, s\})]$$

134
The argument in curly brackets in the exponent can be written as:

\[
\sum_{\{\omega, s\}: \text{s.t. } a_\theta(\omega) = a_\theta(s) = 0} \pi_0 \mu(\{\omega, s\}) + \sum_{\{\omega, s\}: \text{s.t. } a_\theta(\omega) = a_\theta(s) = 1} \pi_1 \mu(\{\omega, s\}) + \sum_{\{\omega, s\}: \text{s.t. } a_\theta(\omega) = 0, a_\theta(s) = 1} \pi_{10} \mu(\{\omega, s\})
\]

and, if we set

\[
\sum_{\{\omega, s\}: \text{s.t. } a_\theta(\omega) = i, a_\theta(s) = j} \mu(\{\omega, s\}) = \psi^\theta_{ij},
\]

then

\[
\psi^\theta_{00} \pi_0 + \psi^\theta_{11} \pi_1 + \psi^\theta_{10} \pi_{10} = \psi^\theta_{00} (\pi_0 - \pi_{10}) + \psi^\theta_{11} (\pi_1 - \pi_{10}) + \pi_{10}
\]

The states \( \theta \) for which \( \lim_{\sigma \to \infty} P_\sigma(\theta) > 0 \) are to be found among those that maximize the above expression on the set \( \psi^\theta_{00} + \psi^\theta_{11} \leq 1 \). Note that the expression is linear in both arguments. If the game is a coordination game (i.e. if \( \pi_{00} > \pi_{10} \) and \( \pi_{11} > \pi_{10} \)), the state that is selected is that where all players adopt the action that corresponds to the risk and (given that the game is balanced) also Pareto dominant equilibrium. If the game has a dominant strategy (for example, if \( \pi_{00} > \pi_{10} > \pi_{11} \)), then the dynamics will select the state where all players adopt the dominant strategy. If the game has two pure strategy Nash-equilibria off the main diagonal (i.e. if \( \pi_{00} < \pi_{10} \) and \( \pi_{10} > \pi_{11} \)) then the dynamics will select the state for which \( \psi^\theta_{00} = \psi^\theta_{11} = 0 \). Analogous considerations apply to a non balanced underlying game, for which, in the same fashion, we obtain that the states with non vanishing limit probability are those that maximize the following expression:

\[
\psi^\theta_{00} (\pi_0 + \pi_{01}) + \psi^\theta_{11} (\pi_{11} + \pi_{10}) + (\pi_{10} + \pi_{01}) (1 - \psi^\theta_{00} - \psi^\theta_{11}) = \psi^\theta_{00} (\pi_0 - \pi_{10}) + \psi^\theta_{11} (\pi_{11} - \pi_{01}) + (\pi_{01} + \pi_{10})
\]
Hence, in a game like the Stag-Hunt game (for which the equilibrium that is risk-dominant is Pareto dominated, because for example $\pi_{00} - \pi_{10} > \pi_{11} - \pi_{01}$ and $\pi_{11} > \pi_{00}$), the equilibrium that will be selected is the equilibrium that is risk-dominant.
REFERENCES


ANDRÁSFAI, B. (1977), Introductory Graph Theory, Bristol: Adam Hilger.


137


138


