Cooperation and free-riding with moral costs

Mirta B. Gordon, Denis Phan, Roger Waldeck, Jean-Pierre Nadal

COOPERATION AND FREE-RIDING WITH MORAL COST

Mirta B. Gordon[1], Denis Phan[2], Roger Waldeck[3], Jean-Pierre Nadal[4]

[1] Laboratoire Leibniz-IMAG, Grenoble, France
mirta.gordon@imag.fr

[2] CREM UMR CNRS 6211, University of Rennes I, France
denis.phan@univ-rennes1.fr

[3] ENST Bretagne, LUSSI, & ICI U.B.O. Brest, France.
roger.waldeck@enst-bretagne.fr

[4] Laboratoire de Physique Statistique, Ecole Normale Supérieure, Paris, France
nadal@lps.ens.fr

ABSTRACT

We study social organizations with possible coexistence at equilibrium of cooperating individuals and pure consumers (free-riders). We investigate this polymorphic equilibrium using a game-theoretic approach and a statistical physics analysis of a simple model. The agents face a binary decision problem: whether to contribute or not to the public good, through the maximization of an additive utility that has two competing terms, a fixed cost for cooperating and an idiosyncratic moral cost for free-riding proportional to the fraction of cooperators. We study the equilibria regimes of this model. We show that there is a fraction of expected cooperators below which cooperation fails to emerge. Besides the homogeneous stable equilibria (everybody cooperates or everybody free-rides), it exists a solution in which cooperators coexist with free-riders. This polymorphic equilibrium is a consequence of the heterogeneous (idiosyncratic) perceptions of the social reproval by the different individuals. We provide analytic results in the case of a simple distribution of the idiosyncratic moral weights, and discuss them on the basis of concepts of game theory.

INTRODUCTION

Basic evidence on several kinds of social organizations whose members are expected to cooperate to a public good reveals a rough partition between individuals that *cooperate* to the public good and *pure consumers* (also called *free-riders*). This polymorphic configuration seems to be a stable form of organization. We propose a model that exhibits the emergence of such polymorphic equilibria.

From a theoretical point of view, there is a large field of research that tries to explain why cooperation may emerge as a stable behaviour among selfish individuals. It is well known that collective action has the structure of a public good problem, of which the prisoner's dilemma is a special case. Those who do not participate to the collective production of the good are not excluded from its consumption. Olson (1965) and early literature about the prisoner's dilemma predict that individuals have no interest to cooperate in the production of a public good, and behave like free-riders. The literature of experimental economics has tested this conjecture. Stable cooperation is seldom attained in finitely repeated public good games (Ledyard, 1995). Typically, although there may be substantial cooperation in the first periods, the level of cooperation decreases through time.

Fehr and Gachter (2000) compare situations where individuals have opportunities for punishing non-cooperators to situations where punishment is absent. They show that introducing a punishment opportunity for the players leads to an increase in cooperation with achievement of almost full cooperation in the last period. This was particularly true for the partner design of their game, in which the same players interact in all the periods. Another way of enforcing cooperation is through social approval of cooperation or disapproval of defection. Gachter and Fehr (1999) consider the importance of social approval and peer pressure on the individuals' behaviours in a public good setting. Individuals may value approval from their peers to a generous contribution from their part. They may also value negatively their own free-riding behaviour when peers contribute. The important point of their paper is that the feeling of group membership and social approval give rise to a large and significant reduction in free-riding. A social norm of high and stable cooperation needs thus some enforcement in order to emerge. Fehr and Fischbacher (2004) examine whether the individuals are willing to support the cost of enforcement.

The model presented in this paper is related to the important body of literature about collective action, recently reviewed by Ostrom (2000) as well as the sociological literature on *threshold models* (following Granovetter 1978) and *critical mass models* (Schelling 1978). For a survey of the sociological approach see Olivier et al. (1993) and Marwell et al. (2001) among others. We analyze a simple situation that corresponds to the following typical scenario: the members of an organization have to share a task whose realization is beneficial for everybody. Cooperators bear a fixed cost for producing the public good for the community. The surplus of all the individuals increases proportionally to the number of cooperators. Individuals that do not cooperate are punished by cooperators through costless moral disapproval. We assume that the costs that come into play are smaller than the agents' endowments, so that the strategy that consists in resigning is not credible.

We study the regimes of equilibrium in games where the individuals are heterogeneous with respect to the value they assign to social disapproval: the cost experienced by a free-rider, proportional to the fraction of cooperators, is idiosyncratically weighted. We do not consider an explicit model of this cost: it may be either a subjective moral burden, or a true sanction. Our main question concerns the conditions under which a polymorphic community with both free-riders and cooperators may exist in equilibrium, and whether there is some threshold of cooperation level below which such a community cannot be stabilized.

The paper is organized as follows: a first section presents the details of the model, which considers the individual decision by the agents given the social context. The following sections are devoted to the study of the equilibrium regimes in a typical case, including dynamical considerations for a simple expectations learning rule.

BASIC MODEL OF BINARY CHOICES WITH SOCIAL SANCTIONS

The basic economic model analyzed in this paper is a particular case of a more general class of models of three strategies studied in Phan et al. (2005), where in addition to the choice between cooperation and defection, an agent must simultaneously decide whether to participate or not. These are generalizations of earlier models of binary choices with externalities (Durlauf, S. 1997; Nadal, J-P. *et al.* 2003, 2004; Phan, D. *et al.* 2004; Gordon, M.B. *et al.* 2005a).

We consider a population of N agents. Each agent i ($1 \le i \le N$) makes one among the following choices:

$$\begin{aligned} s_i &= 1 \quad \text{(to cooperate)} \\ s_i &= 0 \quad \text{(to free-ride)} \end{aligned} \tag{1}$$

Each agent i have a social network of partners (a subset ν_i of the population, of cardinal N_i) who also have to decide which strategy to choose. We assume that each individual i that chooses to cooperate (s_i=1) bears a fixed cost C_i. Free-riders bear a cost due to social disapproval by their partners. There is thus a social dimension in the model, related to the network of relations of the agent, hereafter called his *neighbourhood*.

The fraction of cooperators in the neighbourhood of i is denoted η_c (to simplify the notation we do not write a supplementary subscript i −as we should, since these quantities depend on i's neighbourhood−), and that of free-riders is 1-η_c. Explicitly:

$$\eta_c = \frac{1}{N_i} \sum_{k \in \nu_i} s_k \tag{2}$$

We assume that each agent has a linear surplus function V_i. It includes a constant term H which is the monetary equivalent of the benefit of belonging to the community (equivalent to the initial endowment in experimental settings). Its value is assumed to be large enough so that the agents can afford the costs of their decisions. Beyond H, each agent has a payoff that depends on the behaviour of the other agents. Thus, the surplus function is

$$V_i(s_i | \eta_c) = H + G\,\eta_c - s_i\,C_i - (1 - s_i)X_i\eta_c \tag{3}$$

where G/N_i is the social payoff earned by agent i thanks to each cooperator in his neighbourhood. A free-rider (s_i = 0) supports instead a cost inflicted by each of his neighbours that cooperate. According to the literature on moral features and norms enforcement (*i.e.* Harsanyi 1977, Rabin 1993, Gächter and Fehr 1999, Fehr and Gächter 2002) this cost may be of different natures: it may represent an unmodeled feature *external* to the agent i, like a real sanction, or merely an *internal* moral disagreement −like guilt− due to the existence of cooperating neighbours. In our model, all the neighbours are assumed to produce the same marginal effect on individual i (if he free-rides), equal to X_i/N_i, where $X_i \ge 0$ is an idiosyncratic weight.

INDIVIDUAL BEST RESPONSE ANALYSIS

The surplus (3) can equivalently be written as

$$\begin{aligned} V_i(s_i = 1 | \eta_c) &= H + G\,\eta_c - C_i \\ V_i(s_i = 0 | \eta_c) &= H + (G - X_i)\,\eta_c \end{aligned} \tag{4}$$

Following Phan (2004), and Phan, Waldeck, Gordon and Nadal (2005), the surplus of an individual playing "against a field" of neighbours with a proportion η_c of cooperators and 1-η_c of defectors in equation (4) can be decomposed into the weighted sum of corresponding payoffs in a pure strategy game. The corresponding two-by-two normal form of the game is presented on Table 1.

	$s_k = 1$	$s_k = 0$
$s_i = 1$	$(H+G-C_i)/N_i$	$(H-C_i)/N_i$
$s_i = 0$	$(H+G-X_i)/N_i$	H/N_i

Table 1: Payoff matrix of agent i in a bilateral game against agent k (player i in rows, player k in columns)

Following Monderer and Shapley (1996) this game belong to the class of (weighted) potential games. That is, best-reply sets and dominance-orderings are unaffected if a constant term is added to a column, and if all the columns are multiplied by a constant. The best reply equivalent of the game is given in Table 2, where the columns correspond to the strategies played by an agent representative of the neighbourhood of agent i, who plays s=1 with frequency η_c and s=0 with frequency 1-η_c. This game belongs to the class of coordination games. The payoffs of coordination on the same strategy allow to calculate the cost of unilateral deviation.

	s = 1 (frequency η_c)	s = 0 (frequency 1- η_c)
$s_i = 1$	$X_i - C_i$	0
$s_i = 0$	0	C_i

Table 2: Equivalent payoff matrix of agent i in a bilateral game against the field

In a population game context this cost is not just the corresponding payoff, as in a pure strategy context, but the sum of two terms: the loss of changing the strategy against the fraction of the population coordinated with the agent minus the payoff earned thanks to the (*ex-ante*) proportion of players playing the other strategy. More specifically, the cost of deviating from $s_i = 1$ is $\eta_c(X_i - C_i)$, but playing $s_i = 0$ allows coordination with the fraction $1 - \eta_c$ of the "field" playing strategy s = 0, so that the actual cost of deviation is thus $\eta_c(X_i - C_i) - (1 - \eta_c)\, C_i = \eta_c X_i - C_i$.

Inspection of Table 2 shows that the strategy $s_i = 1$ is better off than $s_i = 0$ against η_c cooperators if $\eta_c(X_i - C_i) > (1 - \eta_c)\, C_i$, that is:

$$s_i = 1 \succ s_i = 0 \Leftrightarrow \frac{X_i}{C_i} > \frac{1}{\eta_c} \tag{5}$$

The marginal agent with $X_i/C_i = 1/\eta_c$ is indifferent between cooperating and defecting. Notice that although the outcome of the game is independent of the parameters G and H, they are essential in the game setting. G (≥0) determines the relative social benefit due to the existence of cooperators. As already stated, we assume that the value of H is high enough for the payoffs be positive whatever the agents' decisions, to ensure it is worth for all the agents to play the game.

Comparison of the payoff of strategy $s_i = 1$ versus that of $s_i = 0$ on the basis of the parameters C_i and X_i gives raise to the following typology of agents:

1. If $X_i \le C_i$, since $X_i\eta_c \le X_i$, it is clear that $s_i=0$ is a strictly dominant strategy against $s_i=1$ for all values of η_c. Agents satisfying this inequality always defect: they are *intrinsic free-riders*.
2. If $C_i < X_i$, then $s_i=0$ may still be a best response against s=1, but since free-riding depends on the value of η_c -see equation (5)- these agents are not intrinsic free-riders, and in fact at equilibrium a fraction, and eventually all, of them will be able to cooperate.

Let us remark that if all the individuals are of type 1, which arises when the largest idiosyncratic weight in the population, X_{max}, is such that $X_{max} \le C_i$ for all i, then the Nash equilibrium with $s_i = 0$ for all i is unique and the game is a prisoner's dilemma.

The general class of *bilateral* games with $C_i \le X_{min}$ is *coordination games* with two equilibria in pure strategies and a single one in mixed strategies. In the case considered in the present paper, each individual plays *simultaneously* against all his neighbours (the "field"), making the issue more involved.

EQUILIBRIUM PROPERTIES OF LARGE POPULATIONS

In the rest of this paper we make the following simplifying assumptions:

- The cost of cooperation is the same for all the agents, $C_i = C$,
- The idiosyncratic weights X_i are quenched (i.e. constant in time) positive random variables, independently and identically distributed throughout the population, with a probability density function f(X) with finite support ($0 \le X_{min} \le X \le X_{max}$)

- The value of H satisfies $H > \max\{C, X_{max} - G\}$ to ensure that no agent has a negative payoff, whatever the adopted strategy.
- The networks of relations have the same structure (homogenous neighbourhood) for all the agents. Moreover, we assume complete connectivity between agents. Thus, the neighbourhood of an agent i, ν_i in equation (2), is the set of all the individuals in the population but himself ($N_i = N - 1$).

We are interested in the properties of very large populations; thus we consider the limit $N \to \infty$. The fraction of cooperators in the neighbourhood of any agent differs from the total fraction of cooperators in the population by at most 1 / (N − 1), which is negligible in this limit.

Since we consider full connectivity in the limit $N \to \infty$, the fraction of cooperators is an unbiased estimator of the probability of cooperating, over the random variable X_i. Let us define the marginal agent having a value of X_i equal to $X_m(\eta_c) \equiv \frac{C}{\eta_c}$, such as he is indifferent between cooperating or not. Then, the equilibrium state of the population satisfies:

$$\eta_c = \frac{1}{N}\sum_{i=1}^{N} s_i = P(X > X_m(\eta_c)) = \Phi(X_m(\eta_c)) \qquad (6)$$

where $\Phi(X) \equiv \int_X^{\infty} f(X')\,dX'$ is the complementary distribution function of the random variable X. Equation (6) is a fixed point equation for the fraction of cooperators. Since $X_i \geq 0$, a *necessary* condition for having full cooperation is that $C \leq X_{min}$, where X_{min} is the lower bound of the support of f(X). That is: only if the cost of cooperation is smaller than the punishment afforded for free-riding for *all* the individuals in the population, then *full* cooperation may be achieved. Notice however that even in this case, this stable solution may not be reached in actual systems, because η_c=0 is *always* a solution of (6). As soon as some individuals have X_i>C (more precisely, a fraction of individuals of finite measure), we have a polymorphic equilibrium where cooperators and free-riders may coexist. The actual proportion of individuals that cooperate depends on the particular distribution of X. Since C is finite, if the support of f(X) is unbounded, the polymorphic equilibrium is generic. But if the support of f(X) is bounded and $C > X_{max}$, the population is no more polymorphic since only the solution η_c=0 exists, with all the agents have vanishing social utility. Figure 1 presents a qualitative picture of the agents' typology in the case of a bounded support.

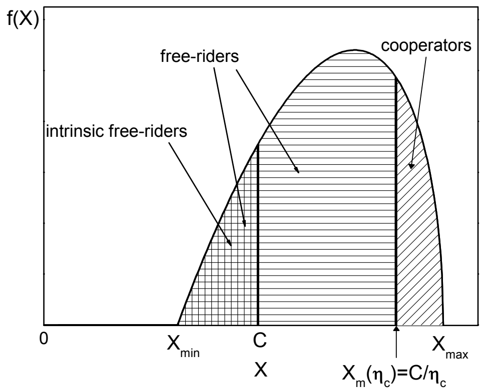

Figure 1. Qualitative typology

It is useful to define reduced variables, $x \equiv X/d$ and $c \equiv C/d$, where d is the width of the distribution f(X). Then equation (6) is equivalent to

$$c = \Gamma(z) \qquad (7)$$

where $z \equiv c/\eta_c$ and:

$$\Gamma(z) = z\,\Phi(z). \qquad (8)$$

UNIFORM DISTRIBUTION

As an example, we consider the case of X drawn from a uniform distribution with support in $[X_{min},X_{max}]$, $d\equiv X_{max}-X_{min}$. Then, $f(X)=1/d$. In terms of the reduced variables,

$$f(x)=\begin{cases}1 & \text{if} \quad x_{min}\le x\le x_{max}\\ 0 & \text{otherwise}\end{cases} \tag{9}$$

and the corresponding cumulative function is:

$$\begin{aligned}\Phi(x)&=\int_x^{x_{max}} f(x')\,dx'\\ &=\begin{cases}0 & \text{if} \quad x_{max}\le x\\ x_{max}-x & \text{if} \quad x_{min}\le x\le x_{max}\\ 1 & \text{if} \quad x\le x_{min}\end{cases}\end{aligned} \tag{10}$$

Then,

$$\Gamma(z)=\begin{cases}0 & \text{if} \quad x_{max}\le z\\ z(x_{max}-z) & \text{if} \quad x_{min}\le z\le x_{max}\\ z & \text{if} \quad z\le x_{min}\end{cases} \tag{11}$$

Figures 2a and 2b represent $\Gamma(z)$ for particular values of x_{min} and x_{max}.

Remembering that $\eta_c=0$ is always a solution, we are left to determine if there exists another solution, $\eta_c=c/z$, where z is given by the intersection between $\Gamma(z)$ and the constant line z=c (see equation (7)). Let us denote

$$c^* \equiv (x_{max}/2)^2 \tag{12}$$

the maximum of the quadratic part of $\Gamma(z)$. We have two possible cases, depending on the position of the maximum of the quadratic part of $\Gamma(z)$:

- $x_{min}\le 1$ (or equivalently, $x_{max}\ge 2x_{min}$) (see Figure 2a). Then
  - if $c>c^*$, only solution $\eta_c=0$ exists.
  - if $x_{min}\le c < c^*$, there are two intersections with $\Gamma(z)$, but only the one corresponding to the smallest value of z is stable, since the corresponding fraction of cooperators *increases* if the (reduced) cost decreases, as it should. The other intersection corresponds to a fraction of cooperators that would *increase* if the cost of cooperating increases. Thus:

$$\eta_c=\sqrt{c^*}\left(1+\sqrt{1-c/c^*}\right) \tag{12}$$

and we have $0\le\eta_c\le\sqrt{c^*}=x_{max}/2$. In this solution the population is composed of a fraction η_c of cooperators *and* a fraction $1-\eta_c$ of free-riders, that is, the population is polymorphic, like in most of the experimental economic settings.

  - if $c\le x_{min}$, the stable solution lies in the linear part of $\Gamma(z)$, so that $\eta_c=1$ is a possible solution.

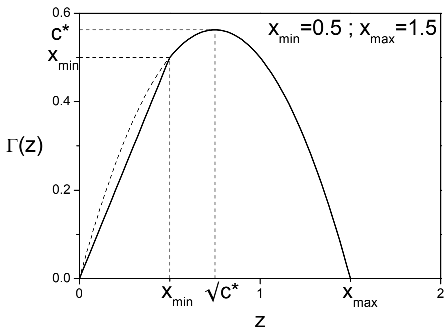

Figure 2a. An example of Γ(z) with $x_{min}<1$.

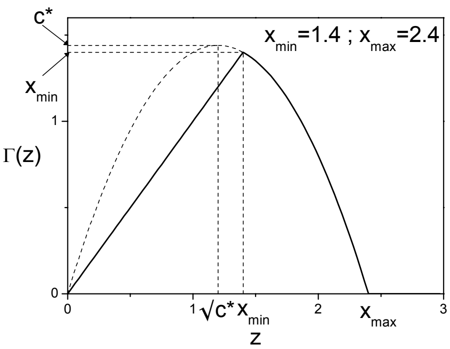

Figure 2b. An example of Γ(z) with $x_{min}>1$.

- $x_{min}>1$ (or equivalently, $x_{max}<2x_{min}$) (see Figure 2b). Then
  - ° if $c>x_{min}$, only the solution $\eta_c=0$ exists.
  - ° if $0<c<x_{min}$, the stable solution lies on the linear part of $\Gamma(z)$, so that the full cooperation solution $\eta_c=1$ also exists.

Since $x_{max}=1+x_{min}$ is *not* an independent variable, the above results may be visualized on a phase diagram in the plane (c,x_{min}), represented on Figure 3, that shows the regions in the parameters space corresponding to the different solutions.

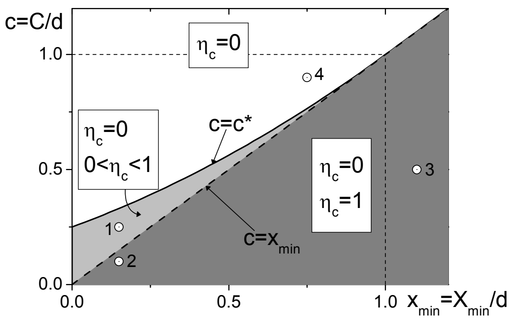

Figure 3. Phase diagram presenting the domains of existence of the different solutions.

Notice that the solution with only free-riders ($\eta_c=0$) exists for all the values of the parameters. In the regions where a solution with finite rates of cooperation exists, it is in competition with the former. Which one will be realized in an actual situation depends on the cognitive properties of the agents. In the simplest case of myopic best response in a repeated game with full information, it depends on the initial guesses by the agents about the expected fraction of cooperators. This situation is typical of coordination games.

If the cost of cooperation vanishes ($C = 0$, so that $c = 0$), there are two possible issues: either all the individuals cooperate, earning $H + G - C$ each, or they are all free-riders, with a payoff $H - X_i$, which may be smaller than the former. The stable equilibria correspond to the entire population selecting the same strategy.

If the cost of cooperation does not vanish, if $C\leq X_{min}$ the situation remains the same as with $C=0$. But if $C>X_{min}$, as far as it remains smaller than $C^* \equiv d\,c^* = \frac{X_{max}^2}{4d}$, the fraction of cooperators, given by equation (12), becomes smaller than 1, and reaches its smallest value at the border $C=C^*$, where it takes the value $\eta_c^* = \frac{X_{max}}{2d}$. This polymorphic solution, with a fraction $\eta_c<1$ of cooperators and $1-\eta_c$ free-riders only exists if $X_{min} < X_{max}/2$. The corresponding parameter values are represented in light grey on figure 3.

Increasing the cost beyond $max\{C^*, X_{min}\}$ hinders cooperation, and the only stable solution is free-riding.

Although the above results are specific of our assumption of a uniform idiosyncratic weights distribution $f(X)$, we expect similar behaviours for more general distributions: if the cost is smaller than the lower bound of the distribution's support, we expect *full cooperation* to be a stable equilibrium in competition with another one corresponding to all the population free-riding. However, if the cost increases, a novel polymorphic solution appears, in which cooperators *coexist* in equilibrium with free riders. The range of parameter values corresponding to this solution depends on the details of the distribution. In the case of a uniform distribution considered here, this phase exists for costs between X_{min} and C^*. The corresponding fraction of cooperators is a decreasing function of the cost, and lies between $\eta_c = 1$ (for $c = X_{min}$) and $\eta_c = X_{max}/2d$ (for $C = C^*$). If the cost is larger than C^*, a distribution-dependent critical value, the solution with 100% of free-riders is the only viable.

DYNAMIC FEATURES

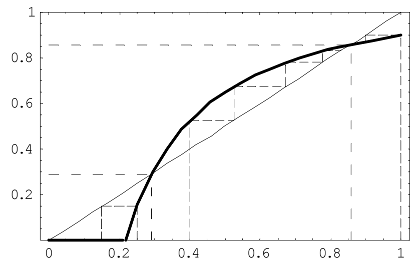

Figure 4.1 X_{min}=0.6, X_{max}=4.6, C=1 (point 1 in fig. 3: x_{min}=0.15,c=0.25) Possible paths are indicated with dashed lines - - -

In the case of a myopic best reply dynamics, individuals are assumed to know the number of cooperators at time t – 1, and play their corresponding best responses at time t. In this case, each individual evaluates the payoffs according to:

$$\begin{aligned} V_i(s_i = 1 | \eta_c(t-1)) &= H + G\ \eta_c(t-1) - C \\ V_i(s_i = 0 | \eta_c(t-1)) &= H + (G - X_i)\ \eta_c(t-1) \end{aligned} \qquad (13)$$

where $\eta_c(t\text{-}1)$ is

$$\eta_c(t-1) = \frac{1}{N}\sum_{k=1}^{N} s_k(t-1) \qquad (14)$$

For large N this gives the following dynamics:

$$\eta_c(t+1) = \Phi(\eta_c(t)) \qquad (15)$$

Depending on the initial conditions (the configuration of strategies at the beginning of the simulations, $\{s_i(t=0),\ 1 \le i \le N\}$), the evolution of the fraction of cooperators may follow different paths: if there is a single solution, it will converge to it, but if there are two possible solutions, like in the grey regions of Figure 3, the reached equilibrium depends on the initial conditions. Figures 4 present $y = \Phi(X_m(\eta))$ given by equation (10), as a function of η (thick curves). The intersections with the diagonals (the lines $y = \eta$) satisfy equation (6) - or its equivalent (7). The broken lines show the learning paths, for the parameter values corresponding to the four points indicated on figure 3. When there are more than one fixed point, paths leading to both fixed points are represented.

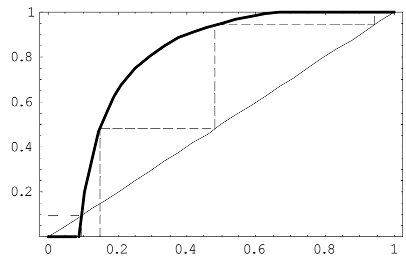

Figure 4.2 X_{min}=1.5, X_{max}=11.5, C=1 (point 2 in fig. 3: x_{min}=0.5,c=0.1)

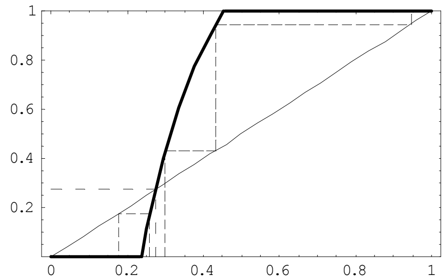

Figure 4.3 X_{min}= 2.2, X_{max}= 4.2, C = 1 (point 3 in fig. 3: x_{min}= 1.1,c = 0.5)

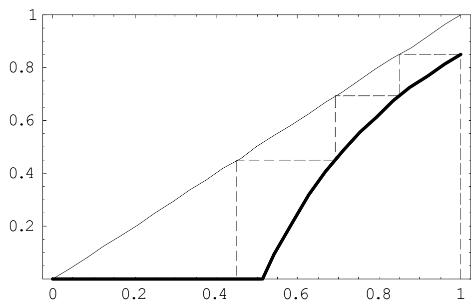

Figure 4.4 X_{min}=0.83, X_{max}=1.94, C=1 (point 4 in fig. 3: x_{min}=0.75,c=0.9)

DISCUSSION

A major interest of population game theory is to provide foundations on equilibrium selection without the assumption of common knowledge (Canning 1995, Blume 1997, Young 1998). Players observe only the strategy of their opponents and have behavioural cognitive capacity (Walliser 1989). That is, players only maximize their payoff according to simple beliefs built on past observations - *i.e.* myopic expectations or more general expectations like Camerer's EWA schemes (Camerer and Ho 1999, Camerer 2003). In our model, each agent plays *against the field*, *i.e.* play the same strategy against all agents in his neighbourhood (Maynard Smith 1982, Crawford, 1990).

The key feature of our discrete choice population game model is that the values X_i are spread over a bounded support $[X_{min},X_{max}]$. As a consequence, bilateral games are asymmetric with respect to the payoffs: all the players have different preferences over the same strategic set. Because agents play against the field, it is nevertheless possible to identify in the phase-diagram some sub-domains where symmetric games concepts and results apply (see also Schmeidler, 1973, and Blonski, 1999). These sub-domains lie in the regions where either $\eta_c=0$ or $\eta_c=1$, because in that cases the cost of deviation, $\eta_c X_i - C$, is either $- C$ or $X_i - C$. Because all the agents play the same strategy, we can analyze the structure of best response despite the heterogeneity of the individual payoffs, and relate some of our results to well known concepts in a symmetric population games

In the game with asymmetric payoffs corresponding to our model, we identify the following parameter configurations (see Figure 5):

1. If $c > x_{max} = x_{min} + 1$ then all agents have $x_i \le c$ and are *intrinsic free-riders*, as already discussed in the typology presented in the section "Individual Best Response Analysis". $s_i = 0$ is a strictly dominant strategy against $s_i = 1$ for all values of η. This zone is located in the north-west of the phase-diagram above the line $c = x_{min} + 1$. The only one equilibrium is $s_i = 0$ for all agents, that is, $\eta_c=1$. The corresponding game is a *prisoner's dilemma*.

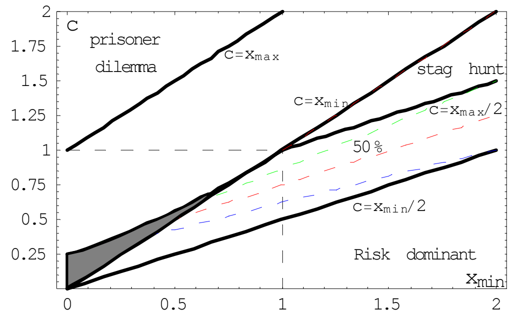

Figure 5. Correspondence with game theory

2. If $c < x_{min}$, the best reply depends on the choice of all the population. This choice is homogeneous, despite the difference in the payoffs. In this zone, all agents have positive values in the diagonal of Table 2. The game belongs to the class of *coordination games*. We obtain two Nash equilibria, $s_i = 0$ for all *i* or $s_i = 1$ for all *i*. It is possible to rank both equilibria according to Pareto dominance and risk dominance (Harsanyi and Selten, 1988). Since all agents have $x_i > c$, the equilibrium $s_i = 1$ Pareto dominates the equilibrium $s_i = 0$ since: $H + G - C > 0$ for all i. Depending on the value of c we identify several regions:
   2.1. Following Harsanyi and Selten (1988), a bilateral equilibrium of coordination between two agents in the cooperation strategy ($s_i = 1$) in Table 2 is said to be *risk dominated* if $c^2 > (x_i - c)(x_j - c)$. This condition is satisfied for all x_i if $x_{min} > c > \sqrt{c^*}$ with c^* given by equation (12), for then $x_{max}/2 < c < x_{min}$, so that $0 < x_i - c < c$ for all the individuals. A unilateral deviation

from the pure strategic equilibrium of coordination on cooperation ($s_i = 1$) is costly for both players, for each couple of players taken at random in the population, and as a consequence for the game against the field. Thus, in this region in the phase diagram the coordination game belongs the class of the so called: "*stag hunt*" game, in which the Pareto dominant equilibrium $\eta_c=1$ is risk dominated by $\eta_c=0$.

2.2. In the sub-region with $c < x_{min} / 2$, the x_i satisfy $x_i - c > c$ for all ì. Coordination on the cooperation strategy ($s_i = 1$) is both Pareto dominant and *risk dominant* since all bilateral equilibria of coordination on cooperation ($s_i = 1$) between any two agents (i,j) satisfy:
$c^2 < (x_i - c)(x_j - c)$ for all x_i, x_j

3. In the intermediate zone: $x_{min} < c \leq x_{max}$, the population of agents is heterogeneous with respect to the dominance structure of their payoff matrix. Some of them are *intrinsic free-riders*, while others as not. The relative weight of *intrinsic free-riders* is given by the relative values of the parameters x_{min} and c (see figure 1). Inside this region with heterogeneous dominance structure across agents, one of the two Nash equilibria is polymorphic: two different strategies (cooperation and free riding) coexist.

CONCLUSION

We analyzed a simple model showing that under very general conditions, polymorphic organizations may exist where cooperators coexist with free-riders. In our model, the individuals have an endowment H large enough to afford the cost due to cooperation or the (moral) cost of free-riding. The latter is proportional to the number of cooperators. Its weight G > 0 sets a lower bound to the value of the individual endowment H for the model being consistent. Notice that the condition for cooperators to exist does not depend on the actual value of G[1]. The latter only affects the payoff, and ensures that the utilities are increasing functions of the degree of cooperation.

We presented our results through the form of a phase diagram, exhibiting the regions in the parameter space where different equilibria are expected. Depending on the parameters these may correspond to full cooperation, pure free-riding, or partial cooperation (the polymorphic equilibria). The existence of the latter is our most interesting result. It is a direct consequence of the heterogeneity in the agent's perceptions, represented by the idiosyncratic weights X_i, of the punishments inflicted by the cooperators to free-riders.

Free-riding is one among possible equilibrium for all the parameter values. Even in the region where full cooperation is stable; its outcome is in competition with pure free-riding. Thus, our model corresponds to a typical coordination game. The actual equilibrium reached by the system depends on the initial conditions. They determine the dynamical paths towards the stationary states of the system.

We are currently investigating how these equilibria are modified if the agents have less information, and evaluate the expected utilities through different learning schemes.

[1] In fact, we are considering in this paper a sub-problem of a more general situation where individuals may choose neither to cooperate nor to free-ride, i.e. to leave the organisation (Phan, Waldeck, Gordon and Nadal 2005).

This work is a part of the project ELICCIR supported by the MNRT / CNRS program "Complex Systems for Human & Social Sciences"

REFERENCES

Arthur, Durlauf, Lane eds. (1997). *The Economy as an Evolving Complex System II*, Santa Fee Institute, Studies on the Sciences of Complexity, Addison-Wesley Pub.Co, Reading Ma,

Axelrod R. (1984). *The Evolution of Co-operation*, Basic Books, New York.

Blonski (1999) Anonymous Games with Binary Actions, *Games and Economic Behavior*, 28 p. 171-180.

Blume L.E. (1997). Population Games, in Arthur, *et al.* (eds.) op. cit., p. 425-460.

Bourgine P., Nadal J.P. eds. (2004). *Cognitive Economy*, Springer Verlag.

Camerer C.F. (2003). *Behavioral Game Theory, experiments in strategic interaction*, Princeton University Press.

Camerer C.F., Ho T. (1999). Experience-weighted attraction learning in normal form games, *Econometrica*, 67 p. 827-874.

Canning D. (1995). Learning and Social Equilibrium in Large Populations, in Kirman, Salmon (eds.) *Learning and Rationality in Economics*, Blackwell, Oxford, p. 308–323.

Crawford V.P. (1990). On the Definition of an Evolutionarily Stable Strategy in the "Plaing the Field" Model, *Journal of Theoretical Biology*, 14 p. 269-273.

Cressman R. (2003). *Evolutionary Dynamics and Extensive Form Games (Economic Learning and Social Evolution)*, The MIT Press, Cambridge, Ma.

Durlauf S.N. (1997). "Statistical Mechanics Approaches to Socioeconomic Behavior", in Arthur, *et al.* op. cit.. p. 81-104.

Fehr E., Gächter S. (2002). Altruistic Punishment in Humans, *Nature* 415, p. 137-140.

Fehr E., Fischbacher U. (2004). Social norms and human cooperation, *Trends in Cognitive Sciences* Vol.8 No.4, April p. 185-190.

Gächter S., Fehr E. (1999). Collective Action as a Social Exchange, *Journal of Economic Behavior and Organization* 39, p.341-369.

Gordon M.B., Nadal J.P., Phan D., Vannimenus J. (2005a). Monopoly Market with externalities: analysis of a simple model, to appears in *Physica A.*

Gordon M.B., Nadal J.P., Phan D. (2005b). "Monopoly Market with externalities: general case", *in preparation.*

Granovetter M. (1978). Threshold Models of Collective Behavior, *American Journal of Sociology*, 83(6) p. 1360-1380.

Harsanyi J.C. (1967/68). Games with Incomplete Information Played by 'Bayesian'Players, Part I, II, and III *Management Science* 14, p. 159-182, 320-334 and 486-502.

Harsanyi J.C. (1977). Morality and the Theory of Rational Behavior, *Social Research*, 44 p. 623-656.

Harsanyi J.C., Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*, The MIT Press, Cambridge, Ma.

Ledyard (1995). Public Goods: A survey of Experimental Research, in Hagel, Roth eds. *Handbook of Experimental Economics* Princeton University Press p.111-194.

Monderer, Shapley, (1996). Potential Games, *Games Economic Behaviour* 14 p.124-143.

Nadal J.P., Phan D., Gordon M.B, Vannimenus J. (2003). Monopoly Market with Externality: an Analysis with Statistical Physics and ACE, *8th Annual Workshop on Economics with Heterogeneous Interacting Agents*, Kiel, May 29-31 (full paper also available at the physicist's base: http://arxiv.org/abs/cond-mat/0311096 and at the economists's WP base:

http://econwpa.wustl.edu/eprints/comp/papers/0312/0312002.abs).

Marwell G., Olivier P.E. (1993). *The Critical Mass in Collective Action, a micro-social theory*, Cambridge University Press.

Maynard Smith J. (1982). *Evolution and the Theory of Games*, Cambridge University Press.

Olivier P.E., Marwell, G. (2001). Whatever Happened to Critical Mass Theory? *A Retrospective and Assessment, Sociological Theory* 19(3) p. 293-311.

Olson M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups,* Harvard University Press, Cambridge, Mass.

Ostrom E. (2000). Collective Action and the Evolution of Social Norms, *Journal of Economic Perspectives*, 14(3) p. 137-158.

Phan (2004). Hierarchy of cognitive interactive agents and statistical mechanics, in Coelho Espinasse Seidel eds., *5th Workshop on Agent Based Simulation,* Lisbon, Portugal, 3-5 May 2004 SCS Publishing House, Erlangen, San Diego, p. 69-76.

Phan D., Nadal J.P., Gordon M.B. (2004). Social interactions in economic theory: an insight from statistical mechanics", in Bourgine, Nadal (eds.), op. cit. p. 225-358.

Phan D., Pajot S. (2005). Complex Behaviours in discrete choice models with social influence, in Aurifeille (ed.), forthcoming, Kluwer.

Phan D., Waldeck R., Gordon M.B., Nadal J.P. (2005). Adoption and Cooperation in Communities: Mixed equilibrium in polymorphic populations, *10th Annual Workshop on Economic Heterogeneous Interacting Agents* (WEHIA 2005), June 13-15, University of Essex, United Kingdom.

Rabin M. (1993). Incorporating Fairness into Game Theory and Economics, *American Economic Review*, 83/5, p.1281-1302.

Schelling T.C. (1973). Hockey Helmets, Concealed Weapons and Daylight Saving A study of binary choices with externalities, *Journal of Conflicts Resolution*, 17-3 September p. 382-428.

Schelling T.C. (1978). *Micromotives and Macrobehavior*, New York, Norton and Company.

Schmeidler (1973). Equilibrium points of non-atomic games *J. Stat. Phys.* 17, p. 295-300.

Walliser B. (1989). Instrumental rationality and cognitive rationality, *Theory and Decision*, 27, p. 7/36.

Young P. (1998) *Individual Strategy and Social Structure*, Princeton University Press.