Physical Biology

PAPER

Punishment in public goods games leads to meta-stable phase transitions and hysteresis

Arend Hintze[1,2] and Christoph Adami[1,2,3]

1. Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI, USA
2. BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI, USA
3. Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA

E-mail: adami@msu.edu

Abstract

The evolution of cooperation has been a perennial problem in evolutionary biology because cooperation can be undermined by selfish cheaters who gain an advantage in the short run, while compromising the long-term viability of the population. Evolutionary game theory has shown that under certain conditions, cooperation nonetheless evolves stably, for example if players have the opportunity to punish cheaters that benefit from a public good yet refuse to pay into the common pool. However, punishment has remained enigmatic because it is costly and difficult to maintain. On the other hand, cooperation emerges naturally in the public goods game if the synergy of the public good (the factor multiplying the public good investment) is sufficiently high. In terms of this synergy parameter, the transition from defection to cooperation can be viewed as a phase transition with the synergy as the critical parameter. We show here that punishment reduces the critical value at which cooperation occurs, but also creates the possibility of meta-stable phase transitions, where populations can 'tunnel' into the cooperating phase below the critical value. At the same time, cooperating populations are unstable even above the critical value, because a group of defectors that are large enough can 'nucleate' such a transition. We study the mean-field theoretical predictions via agent-based simulations of finite populations using an evolutionary approach where the decisions to cooperate or to punish are encoded genetically in terms of evolvable probabilities. We recover the theoretical predictions and demonstrate that the population shows hysteresis, as expected in systems that exhibit super-heating and super-cooling. We conclude that punishment can stabilize populations of cooperators below the critical point, but it is a two-edged sword: it can also stabilize defectors above the critical point.

1. Introduction

When individuals maximize their self-interest by exploiting a public good, they are often doing so by harming their (and others') own long-term interest, and create a social dilemma termed the 'tragedy of the commons' [1]. The tragedy of the commons is often discussed in environmental politics (for example, overgrazing and overfishing), as well as social science and politics (for example, vandalism and taxation) [1]. However, the tragedy of the commons also plays an important role in evolutionary biology [2]: rate-yield tradeoffs in bacterial metabolism [3], the evolution of virulence [4], and the manipulation of a host by a group of parasites [5] can be viewed as social dilemmas involving a public good. Social dilemmas [6] (such as the tragedy of the commons) can be studied within the framework of evolutionary game theory (EGT) [7–12], which describes populations of agents engaging in pairwise (or groupwise) interactions, with defined payoffs for different strategies. The tragedy of the commons is usually described by a particular game form known as the 'public goods' game.

The public goods game is a standard within the field of experimental economics [13–15]. In this game, players possess tokens that they can invest into a common pool (the public good). The total sum contributed by the players is then multiplied by a 'synergy

factor' thus creating a positive yield. This amount (typically larger than the invested sum) is then equally distributed to the players in the pool, irrespective of whether they invested or not. A group of players maximizes their investment if all the players contribute (so as to take maximum advantage of the synergy). However, this behavior is vulnerable to 'free-riders': individuals that share in the pool but do not invest themselves. It can easily be shown that the rational Nash equilibrium for this game is *not* to pay in, because this strategy clearly dominates all others regardless of their play [1].

Hardin originally suggested that the tragedy of the commons can only be avoided by *punishing* free riders [1]. Indeed, it has been shown that punishment can counteract defectors effectively [16–26], but punishment is difficult to maintain because it is costly and may reduce the mean payoffs of group members compared to groups in which punishment of free-riders is not possible [27].

It was previously thought that punishment cannot be maintained in well-mixed populations [28] so most of the literature has focused on the spatial version of the game, where analytical results are difficult (but not impossible) to obtain [29]. More recent work has uncovered a decidedly more complicated picture, in particular because punishment can be performed in a number of different ways, that each change the dynamics of the population considerably. For example, if the punisher also rewards cooperators at the same time [30], complex dynamics that depend on the strength of punishment and the size of the synergy can emerge in the spatial game; however the reward/punish strategy is only stable in rare circumstances. It is also possible to punish in such a manner that the severity of punishment depends on the number of defectors [31]. In the spatial version of the game, introducing such an 'adaptive' punishment leads to a strong enhancement of the cooperative phase, as rare punishers protect the boundary between the phases with maximum punishment. Another variation of punishment is to introduce a conditional punisher, that scales punishment with the number of defectors in the group [32]. In that game (also in the spatial setting), conditional and unconditional punishers cannot coexist, and who ultimately wins depends on the strength of punishment. There are variations of the public goods game in which punishment can be maintained in other ways, for example by voluntary punishment [29, 33–36] or by using a pool (that is institutionalized) as opposed to peer punishment [36–39], but we do not study those here.

A variant of the game that is closer to the one we study here involves *probabilistic* punishment, where rather than punishing any defector with certainty, cooperators punish with probability π. Chen *et al* [40] show that introducing a probability to punish changes the fixed point structure of the game in the well-mixed

regime so that a repulsive interior fixed point can appear, and this enhances cooperation also in the spatial version of the game.

Here we study the well-mixed version of the public goods game with probabilistic punishment and probabilistic cooperation, establish a number of theoretical results that suggest complex dynamics at the interface between the cooperating and defecting phases (as a function of the synergy factor) and clarify the role of punishment as a catalyst of cooperation in extensive agent-based simulations that agree with the theoretical results.

2. Mean field theory of public goods games

The public goods game emulates strategic decision making by groups, in which an individual must select between different decisions that affect the group as a whole. Each individual in a group of $k + 1$ players (the focal player and her k participants) can decide to cooperate by making a contribution of 1 unit to the public good, while defecting individuals do not contribute.

The sum of all contributions from cooperating players is multiplied by r (the synergy factor) and divided among all players. If N_C is the number of cooperators within the group (but not counting the focal player, i.e., $N_C \leqslant k$) and N_D is the number of defectors, then the cooperator obtains a payoff

$$P_C = r\frac{(N_C + 1)}{k + 1} - 1 \qquad (1)$$

compared to the defector's

$$P_D = r\frac{N_C}{k + 1} \ . \qquad (2)$$

A dilemma exists if it is advantageous for the individual to defect, while mutual cooperation would be best for all. Clearly a defector does better if $P_D - P_C > 0$, so a dilemma exists only if $r < k + 1$. At the same time, the payoff for a cooperator playing within a group of cooperators should be larger than the payoff for a defector playing only with defectors, that is, $P_C(N_C = k) - P_D(N_C = 0) > 0$ which implies $r > 1$. Thus, a dilemma exists only for $1 < r < k + 1$ (see figure 1). Standard EGT arguments imply that defection is the rational (and optimal) strategy for $r < k + 1$, while cooperation is selected for when $r > k + 1$. The synergy parameter r can thus be viewed as a critical parameter, dialing a phase transition from defection to cooperation as r is increased through $k + 1$. This transition is (in the limit of infinite population size and vanishing mutation rate) of first-order (meaning an abrupt discontinuous transition from defection to cooperation as the synergy parameter is raised across its critical value), as can be seen from the replicator equation for the density of cooperators ρ_C
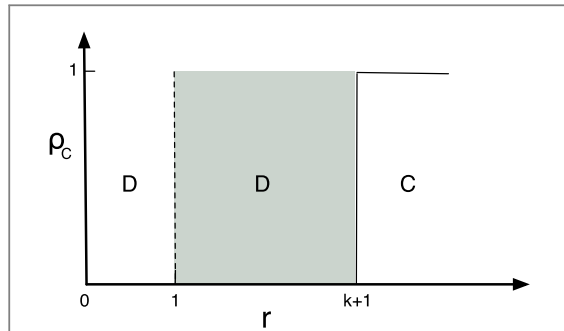
Figure 1. The phase diagram of the public goods game with a synergy factor r. Below $r = 1$ defection is the strategy with the highest payoff and therefore favored by evolution. Conversely, above $r = k + 1$ cooperation is evolutionarily favored. A dilemma exists in the gray-shaded area between $r = 1$ and $r = k + 1$, where cooperation would be beneficial for a cooperating group as a whole, but defection is the Nash equilibrium point and thus evolutionarily favored.
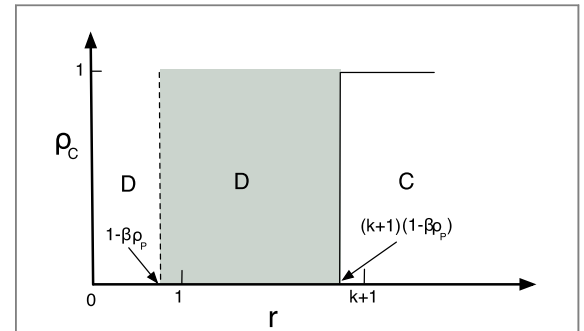

Figure 2. The phase diagram of the public goods game with a synergy factor r and punishment. Compared to figure 1, the area where a dilemma occurs is shifted toward lower r, implying that cooperation can occur for smaller r (right boundary of dilemma area). A single defector cannot invade the population to the right of the critical point, and a single cooperator cannot invade to the left of r_c.

$$\dot{\rho}_C = \rho_C \left(1 - \rho_C \right) \left(P_C - P_D \right). \tag{3}$$

Note that the public goods game turns into the standard Prisoner's dilemma for $k = 1$, with a dilemma for $1 < r < 2$.

How can this dilemma be solved? How can evolution achieve cooperation in the gray-shaded area in figure 1? The answer is: this is impossible unless additional mechanisms change the critical point below $r_c = k + 1$. One such mechanism investigated in the literature is giving players the option to punish players who do not contribute. Following the notation of Helbing *et al* [26], defecting players suffer a *fine* β/k levied by each punisher in the group, which costs each punisher a *penalty* γ/k. Let N_M be the number of players that cooperate as well as punish (the 'moralists') and N_I the number of defectors that punish ('immoralists'). As before, N_C and N_D are the number of players that cooperate, namely defect but do not punish. The payoffs for the four possible strategies then become

$$P_C = r \frac{\left(N_C + N_M + 1 \right)}{k + 1} - 1, \tag{4}$$

$$P_D = r \frac{\left(N_C + N_M \right)}{k + 1} - \beta \frac{\left(N_D + N_I \right)}{k}, \tag{5}$$

$$P_M = P_C - \gamma \frac{\left(N_D + N_I \right)}{k}, \tag{6}$$

$$P_I = P_D - \gamma \frac{\left(N_D + N_I \right)}{k}. \tag{7}$$

2.1. Mean-field theory

Let us calculate the critical point for the game with punishment, assuming a well-mixed population so that each player encounters on average the same fraction of strategies. Introducing the mean density of cooperators ρ_C and the mean density of punishers ρ_P

$$\rho_C = \frac{N_C + N_M}{k}, \tag{8}$$

$$\rho_P = \frac{N_M + N_I}{k}, \tag{9}$$

along with $\rho_D = \dfrac{N_D + N_I}{k}$ we can write the average payoffs for each of the four strategies as

$$P_C = r \frac{\left(k\rho_C + 1 \right)}{k + 1} - 1, \tag{10}$$

$$P_D = r \frac{k\rho_C}{k + 1} - \beta \rho_P, \tag{11}$$

$$P_M = P_C - \gamma \rho_D, \tag{12}$$

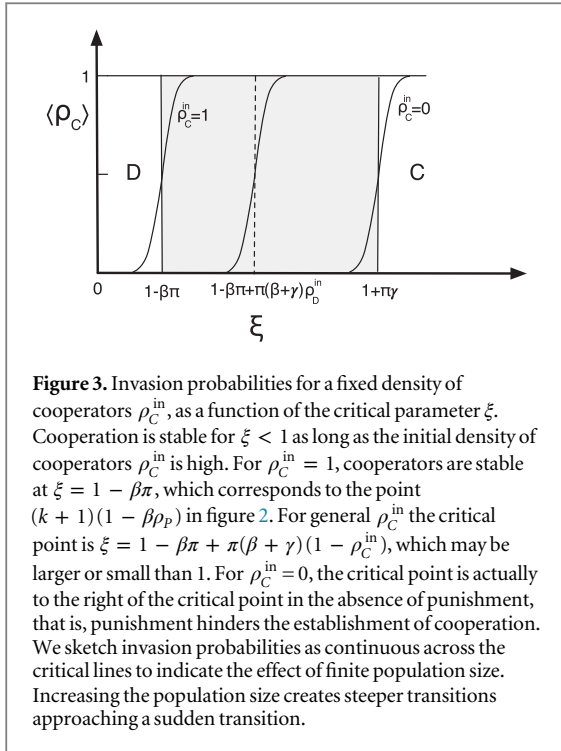$$P_I = P_D - \gamma \rho_D. \tag{13}$$

Investigating $P_C - P_D$ again, we notice that the area where the dilemma exists is now shifted by $\beta\rho_P$ (see also figure 2):

$$1 - \beta\rho_P < r < (k + 1)\left(1 - \beta\rho_P \right), \tag{14}$$

where the right boundary corresponds to the critical point $r_c = (k + 1)(1 - \beta\rho_P)$ separating a cooperating and a defecting phase. Because we will be concerned with this critical point from now on, let us introduce the re-scaled synergy parameter $\xi = r/(k + 1)$. The critical point is then $\xi_c = 1 - \beta\rho_P$. According to standard population genetics, a single cooperating individual cannot invade a population of defectors unless its fitness advantage $P_C - P_D$ is positive, which implies $\xi > 1 - \beta\rho_P$. However, if the entire population consists of defectors, punishment is expected to be absent because defectors do not punish each other (i.e., we assume here that immoralists do not matter in the long run, as was found in numerical simulations [26]).

What happens if a *group* of cooperators (rather than a single individual) tries to invade the defectors (or a group of defectors tries to invade the cooperators)? Because the fitness of any group is frequency-dependent, we have to recalculate the mean fitness of a group as follows: assume a population of strategies given by the mean densities (ρ_C, ρ_P). Let us also assume that, in general, cooperators punish with a probability π_C, while defectors punish with a

Figure 3. Invasion probabilities for a fixed density of cooperators ρ_C^{in}, as a function of the critical parameter ξ. Cooperation is stable for $\xi < 1$ as long as the initial density of cooperators ρ_C^{in} is high. For $\rho_C^{\text{in}} = 1$, cooperators are stable at $\xi = 1 - \beta\pi$, which corresponds to the point $(k + 1)(1 - \beta\rho_P)$ in figure 2. For general ρ_C^{in} the critical point is $\xi = 1 - \beta\pi + \pi(\beta + \gamma)(1 - \rho_C^{\text{in}})$, which may be larger or small than 1. For $\rho_C^{\text{in}} = 0$, the critical point is actually to the right of the critical point in the absence of punishment, that is, punishment hinders the establishment of cooperation. We sketch invasion probabilities as continuous across the critical lines to indicate the effect of finite population size. Increasing the population size creates steeper transitions approaching a sudden transition.

probability π_D. Then, $\rho_P = \pi_C \rho_C + \pi_D \rho_D$. The mean fitness of a group of cooperators is then given by

$$\bar{w}_C = (1 - \pi_C)P_C + \pi P_M = P_C - \gamma\pi_C\rho_D, \quad (15)$$

using the payoffs (10)–(13), and the fitness advantage of the cooperating type with respect to the defectors is

$$\begin{aligned} \bar{w}_C - \bar{w}_D &= P_C - P_D - (\pi_C - \pi_D)\gamma\rho_D \\ &= \xi - 1 + \beta(\pi_C + \pi_D) \\ &\quad - (\pi_C - \pi_D)(\beta + \gamma)\rho_D. \end{aligned} \quad (16)$$

We will see in the numerical results below that immoralists go extinct quickly (because they bear the double cost of meting out and receiving punishment). As a consequence, we set $\pi_D = 0$ (defectors do not punish), and write the cooperator's probability to punish as $\pi_C \equiv \pi$, so that

$$\bar{w}_C - \bar{w}_D = \xi - 1 + \beta\pi - \pi(\beta + \gamma)\rho_D. \quad (17)$$

Equation (17) implies that punishment enables a 'premature' phase transition to cooperation as long as a 'nucleus' of cooperators ρ_C^{in} of sufficient size exists: a hallmark of metastability (see figure 3). Thus, a 'fluctuation' of pure moralists ($\rho_C^{\text{in}} = 1$, $\pi = 1$) is stable at $\xi = 1 - \beta$, which can be significantly smaller than 1 if the effect of punishment is large. However, the opposite dynamics occur for groups of defectors: they can invade stable cooperators at $\xi > 1$ as long as the density of invading defectors ρ_D^{in} is large enough. As outlined in figure 3, a 'fluctuation' into all defectors $\rho_D^{\text{in}} = 1$ is stable for $\xi = 1 + \pi\gamma$, which can be substantially larger than 1 when the defectors displace perfect moralists ($\pi = 1$). Thus, punishment enables both cooperation and defection in a meta-stable phase, away from the critical point $\xi = 1$. This behavior is akin to the phenomenon of supercooling/

superheating, and can result in hysteresis: a population that starts in a defecting phase will stay in the cooperating phase past the critical point $\xi = 1$ as ξ is raised adiabatically from low values, and remain in the defecting phase past the critical point as ξ is lowered from high values adiabatically. We will verify this behavior in the numerical simulations that follow.

3. Evolutionary simulation of public goods games

In this section we test the predictions of the (infinite population size) mean-field theory using agent-based simulations with finite population size. The population consists of 1024 individuals who each have four (randomly assigned) opponents, that is, we use $k = 4$ throughout in the results presented here (with some results for $k = 8$). For populations of this size, neutral drift is negligible and results do not change qualitatively if populations are larger. However, the steepness of the transition between defection and cooperation may depend on the population size in the standard manner expected from finite-size scaling arguments (see, e.g., [41, p 441]).

3.1. Game dynamics and genetic algorithm

Since all opponents are also players, each individual plays $k + 1$ games per update. The actual play of each individual is determined by their probabilities to cooperate p_C and to punish p_P encoded as two genetic loci, which can be thought of as the outcome of a network of genes that encode this decision. When mutating strategies, instead of mutating the individual genes that make up the decision pathway, we simply replace the parental probability p_C by a uniformly drawn random number in the offspring. We will call the locus encoding the probability p_C simply the 'C gene' and similarly for the punishment gene.

When every individual has played against its k partners, 2% of the population is replaced using a Moran-like process [42] in a well-mixed fashion. The Moran-like process with a finite replacement rate interpolates between a true Moran process (replacement rate equal to inverse population size) and a Wright–Fisher process, where the entire population is replaced every update. In our replacement scheme, the identity of the players in any group is unrelated to their ancestry so that, effectively, the members of a particular playing group are randomly selected from the population [43]. With a replacement rate of 2%, it takes on average 50 population updates until the entire population is replaced, that is, a single generation has elapsed. In our simulations, the fitness of each individual is cumulative, that is, the payoff obtained in the next update of the population is added to the payoff already obtained (until that player is removed). However, we have tested that zeroing out the fitness after each update does not alter the game dynamics. We also

verified that varying the replacement rate does not change the dynamics of the population in this game, unlike in the case where strategies communicate [44]. Indeed, if strategies make their play dependent on the last play, then replacing the opponent can introduce noise into the communication, resulting in different levels of cooperation.

We verified that the probability for a player to encounter cooperators is independent of whether that player is a cooperator or a defector, as is required for well-mixed populations [45]. The accumulated payoff (fitness) is used to calculate the probability that this player's strategy will be chosen to replicate and fill the spot of a player that was removed in the Moran process. In case payoffs (calculated according to the equations above) are negative, we add a constant payoff to each and every strategy so that the relative payoffs are unchanged (it is known that such an offset does not alter the population dynamics). While the spatial version of the game shows somewhat different dynamics than studied here, we study the well-mixed version because it is amenable to theoretical prediction (see section 2).

The two genes of every individual mutate with a probability μ when replicated. As mentioned earlier, mutating a probability replaces the probability with a uniformly distributed random number. We used a fixed mutation probability ($\mu = 2\%$ per locus) in the results presented here, except when we test the influence of the mutation rate on the broadening of the phase transition. We have previously studied the effect of varying mutation rate in this game [46] and found only a weak dependence on the location of the fixed point.

3.2. Line of descent (LOD)
After 500 000 updates, the LOD of the population is reconstructed [47, 48], by picking a random organism of the final population and following its ancestry all the way back to the starting organism. This is possible because no recombination occurs between genotypes: descent is entirely asexual. The LOD recapitulates the evolutionary dynamic of the population, because it contains the successive list of genotypes that have achieved fixation in the population. Because the population size is large, only a small fraction of mutations (on the order $1/N$ where N is the population size) find themselves on the LOD by chance. Thus, the LOD reflects the selective pressures operating on the population, and the fixed point of the evolutionary trajectory faithfully characterizes these pressures. The ancestral genotype that anchors all lines of descent is given by the random strategy $p_C = 0.5$ and $p_P = 0.5$. Because there is only one species in these populations, the individual LODs of the population coalesce to a single LOD fairly rapidly (which is why it is sufficient to pick a random genotype for following the LOD). In other words, the common ancestor of the entire
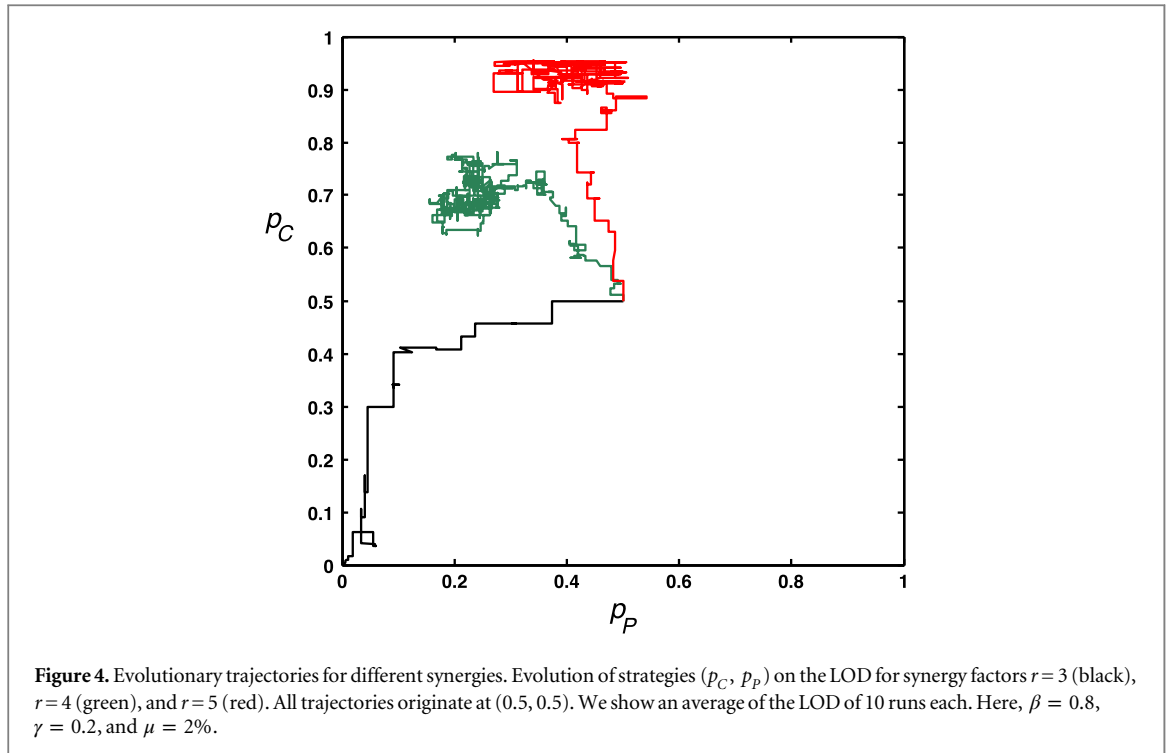
population is invariably fairly recent. To be certain that we deal with LODs that have coalesced when calculating strategy fixed points from the LOD, we routinely discard the last 50 000 updates (about 1000 generations) from every run. When determining evolutionary fixed points for the trajectory, we also discard the first half, as the population trajectory may still be transient.

4. Results

4.1. Evolutionary trajectories and fixed points
As the strategies adapt to the environmental conditions (specified by the parameters that define the game, including the neighborhood size, the mutation rate, and the replacement rate), the probabilities change from their initial values $(p_C, p_P) = (0.5, 0.5)$ toward the selected 'fixed point' strategy. In order to visualize the evolutionary trajectory of a population, we reconstruct the evolutionary LOD of an experiment (LOD, see section 3.2), which tells the story of that adaptation, mutation by mutation. While the LOD in each particular run can show probabilities varying wildly, averaging many such LODs can tell us about the selective pressures the populations face. In particular, averaging the probabilities on the LODs after they have settled down, can tell us the *fixed point* of evolutionary adaptation [44]. We determine this fixed point by discarding the first 250 000 updates of every run (the transient), along with the last 50 000 (in order to remove the dependence of the LOD on the randomly chosen anchor genotype) and averaging the remaining 200 000 updates. Note that this fixed point is a computational fixed point only: we do not mean to imply that the population's genotypes all end up on this exact point. Rather, due to the nature of the game and the selective pressures that change as the composition of the population changes, the evolutionary trajectories approach this point and then fluctuate around or near it. Thus, the fixed point reflects the *mean* successful strategy given the conditions of the game.

We show in figure 4 the average trajectories for three different synergy factors $r = 3$, 4, and 5 all anchored at the random strategy $(p_C, p_P) = (0.5, 0.5)$ that was used as the seed strategy for every evolutionary run. We can see that, depending on the synergy (and the values chosen for the cost and effect of punishment), populations evolve toward a cooperating or defecting fixed point, and take different trajectories to get there. For $r = 3$, synergy is too low to lead to cooperation, and the fixed point of that trajectory is $(p_C, p_P) = (0, 0)$, that is, defection. For $r = 4$, however, the population moves toward a fixed point centered around $(p_C, p_P) = (0.7, 0.2)$, that is, players cooperate most of the time. (The location of the endpoint of the trajectory does not depend on the starting point.) Note, however, that the players engage in

Figure 4. Evolutionary trajectories for different synergies. Evolution of strategies (p_C, p_P) on the LOD for synergy factors $r = 3$ (black), $r = 4$ (green), and $r = 5$ (red). All trajectories originate at (0.5, 0.5). We show an average of the LOD of 10 runs each. Here, $\beta = 0.8$, $\gamma = 0.2$, and $\mu = 2\%$.

punishment only sparingly. For $r = 5$, cooperation is almost fully established, while punishment occurs about 40% of the time on average. However, the average trajectory (average over ten independent runs) only tells part of the story, because at this level of cooperation there is very little difference between a punishing and a non-punishing player (given there are very few players to punish) and as a consequence the punishment gene has begun to drift. An unselected (and thus drifting) probability p_P is a uniformly distributed random number, with mean 1/2 and variance 1/12. As $p_C \to 1$, the average p_P and its variance approach precisely these numbers.

When mapping the strategy fixed point (average strategy on the LOD over 20 independent runs, again discarding the transient and the last 50 000) as a function of the parameters β (effectiveness) and γ (cost) of punishment (defined in section 2) each in the range from 0.0 to 1.0 and at low synergy $r = 3.0$, we find that defection is the most prevalent strategy on the LOD (see figure 5(A)), as was found previously [25, 26]. When $\gamma = 0$ there is no cost associated with the punishment, which implies that the P gene is not under selection and drifts. Thus, for this value of synergy (and lower), we find that the strategy fixed point is defection without punishment, except for the values $\gamma = 0$, where punishment is random.

As the degree of synergy increases to $r = 3.5$, cooperation starts to appear even in this well-mixed population (see figure 5(B)), while it appears as early as $r = 2$ for sufficiently high β and low γ in the spatial (but deterministic) version of the game, see [25, 26]. For $r = 4$ we find players cooperating ($p_C \approx 0.8$) at high β and low γ which indicates that under conditions where

punishment is not very costly or even free, punishment pays off. In addition we notice that the probability to punish increases under the same conditions that allows cooperation (high β and low γ, that is high impact, low cost of punishment), indicating that punishment is indeed used to enforce cooperation (figure 5(C)). The mean punishment probability grows to 0.5, but at the same time the variance shows that this gene is not under selection (as long as $\gamma \neq 0$).

Increasing the synergy level even more toward $r = 4.5$ we witness the emergence of dominance of cooperation ($p_C > 0.5$) for most of the range of punishment cost and effectiveness, see figure 5(D). At the same time the punishment probability reaches 0.5 for a larger range of parameters, but the mean punishment probability on the LOD never exceeds 0.5, implying that full persistent punishment is not stable, and probably not necessary. Note that, in an implementation where decisions are deterministic (such as in the implementation of Helbing *et al* [26]), punishment may remain for a long time in the population even though it is not selected anymore. In that case, players that cooperate with and without punishment have exactly the same fitness, and one or the other strategy should only dominate by drifting to fixation neutrally, a process that can take a significant amount of time in large populations such as those studied in [26].

4.2. Critical dynamics and the role of punishment
Previously, a phase transition between cooperative and defective behavior in the public goods game as a function of the synergy r was observed for the spatial version [25, 28, 29] of the game (but not the well-
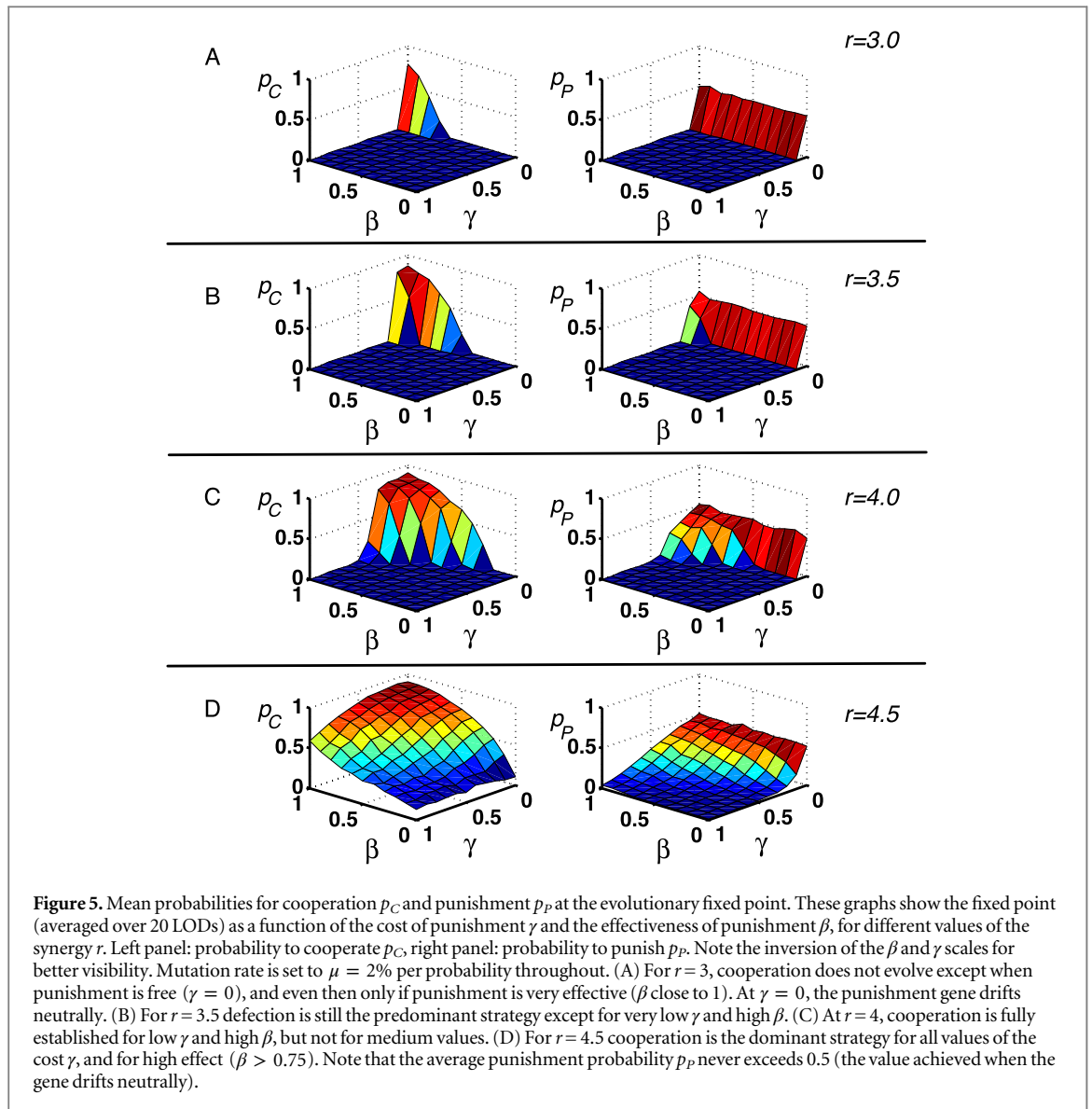
Figure 5. Mean probabilities for cooperation p_C and punishment p_P at the evolutionary fixed point. These graphs show the fixed point (averaged over 20 LODs) as a function of the cost of punishment γ and the effectiveness of punishment β, for different values of the synergy r. Left panel: probability to cooperate p_C, right panel: probability to punish p_P. Note the inversion of the β and γ scales for better visibility. Mutation rate is set to $\mu = 2\%$ per probability throughout. (A) For $r = 3$, cooperation does not evolve except when punishment is free ($\gamma = 0$), and even then only if punishment is very effective (β close to 1). At $\gamma = 0$, the punishment gene drifts neutrally. (B) For $r = 3.5$ defection is still the predominant strategy except for very low γ and high β. (C) At $r = 4$, cooperation is fully established for low γ and high β, but not for medium values. (D) For $r = 4.5$ cooperation is the dominant strategy for all values of the cost γ, and for high effect ($\beta > 0.75$). Note that the average punishment probability p_P never exceeds 0.5 (the value achieved when the gene drifts neutrally).
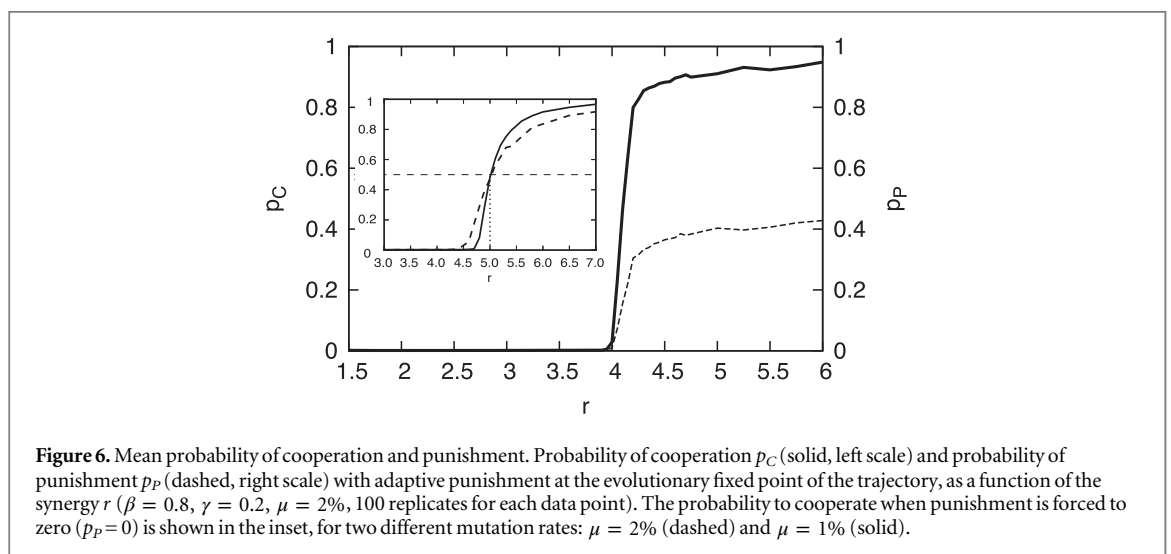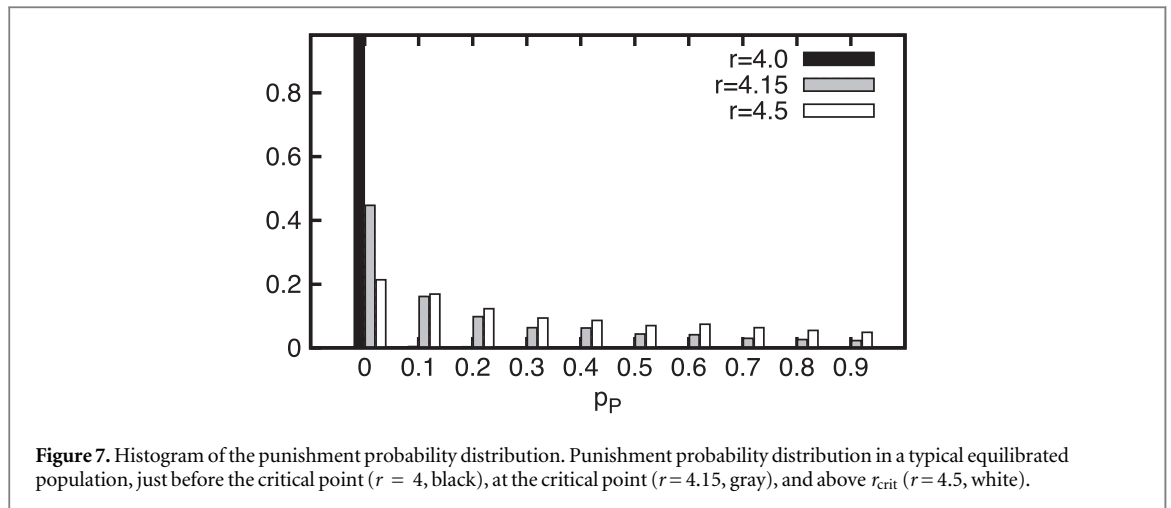

Figure 6. Mean probability of cooperation and punishment. Probability of cooperation p_C (solid, left scale) and probability of punishment p_P (dashed, right scale) with adaptive punishment at the evolutionary fixed point of the trajectory, as a function of the synergy r ($\beta = 0.8$, $\gamma = 0.2$, $\mu = 2\%$, 100 replicates for each data point). The probability to cooperate when punishment is forced to zero ($p_P = 0$) is shown in the inset, for two different mutation rates: $\mu = 2\%$ (dashed) and $\mu = 1\%$ (solid).

mixed version). We can study the critical point and its dependence on punishment in detail in the well-mixed version of the game, where analytical predictions (as outlined above) are available. We show in figure 6 the average probability to cooperate (solid line) and to punish (dashed line) as a function of synergy for our

Figure 7. Histogram of the punishment probability distribution. Punishment probability distribution in a typical equilibrated population, just before the critical point ($r = 4$, black), at the critical point ($r = 4.15$, gray), and above r_{crit} ($r = 4.5$, white).

default values $\gamma = 0.2$ and $\beta = 0.8$. Cooperation sets in at $r = 4$ and becomes prevalent for synergies just exceeding that. The simulations show that the phase transition is broadened from the expected first-order behavior, owing to two factors: finite population size and finite mutation rate. The inset in figure 6 shows the transition in the absence of punishment at two different mutation rates, suggesting that mutations introduce 'disorder-broadening' [49], which can lead to complete 'rounding' of the discontinuous transition.

We will now study how punishment affects the critical point. The average probability of cooperation in figure 6 shows the typical behavior of an order parameter as a function of the critical parameter r in a broadened first-order transition. Thus, although punishment is sporadic when it is possible—and drifts when cooperation is established—it is essential to lower the critical barrier for cooperation. The probability distribution of the punishment gene throughout the population (figure 7) shows that punishment is never prevalent: it is absent below the critical point, while the distribution is close to uniform (because of drift) above it. In a sense, punishment catalyzes the transition from defection to cooperation. Note also that the levels of cooperation achieved (at a given r) are significantly higher when punishment exists, even though punishment is only weakly selected for. Apparently, the possibility of punishment alone is sufficient to enforce higher levels of cooperation, but the mechanism for this enforcement is not immediately clear as punishment is rare above the critical point.
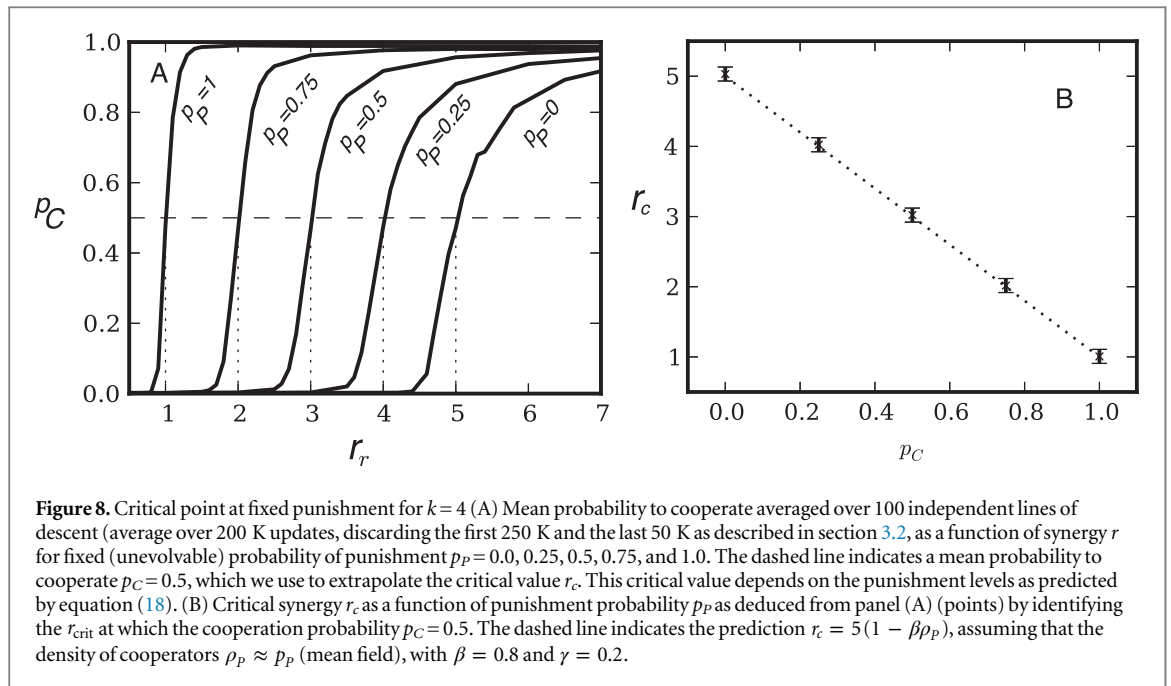
In section 2 we calculated approximately the point at which cooperation is favored in a mean-field approach that does not take mutations into account, by writing equations (4)–(5) in terms of the density of cooperators ρ_C encountered by players in a group, and found that cooperation was favored as long as

$$r > (k + 1)\left(1 - \beta\rho_P\right). \qquad (18)$$

This equation (which also follows from a replicator equation approach for the corresponding deterministic game) implies that the emergence of cooperation

depends crucially on the density of punishers. In fact, the mean-field theory predicts that cooperation in the absence of punishment is favored only at $r = 5$. We see cooperation emerge quite a bit earlier than in our simulations (see inset in figure 6), but crosses $p_C = 0.5$ very close to $r = 5$, as predicted by the mean field theory. Of course, the departure from the mean-field theory results is a consequence of the finite population size and mutation rate of the simulations. We also note that while the simulations suggest a stable fixed point $p_C < 1$ above the critical point, this does not mean that the corresponding fixed point for the deterministic game admits a stable mixture of cooperating and defecting strategies. While from general arguments the fixed point of the stochastic game must be given by the corresponding fixed point of the deterministic game [50] (see also [51]), the 'broadening' of the phase transition via mutations (the equivalent of random quenched impurities [49]) leads to a stochastic fixed point with $p_C < 1$. But note that this fixed point is statistical only: for each run the trajectory fluctuates around this point.

We can test equation (18) explicitly by finding the critical r at which p_C crosses 0.5 for simulations in which the punishment probability is held fixed, so that $\rho_P \approx p_P$. To find the critical point, we performed 100 simulations each at fixed r with small increments Δr and interpolated the data within the steep portion of the transition to find the crossover point. The curves in figure 8 show that the steepness of the transition between cooperation and defection depends on the level of punishment, changing from a dependence reminiscent of second-order transitions (at vanishing punishment) toward a first-order-like transition at high punishment. We plot the critical line $r_c = (k + 1)(1 - \beta\rho_P)$ in figure 8 for $k = 4$ and $\beta = 0.8$ ($r_c = 5 - 4p_P$). The mean field theory reproduces the simulated r_c within errors. The prediction in fact works just as well for other parameter values: we tested $k = 8$ (each agent plays with eight random other agents) and readily observe that the critical value is given by $r_c = 10 - 8p_P$ (data not shown).

Figure 8. Critical point at fixed punishment for $k = 4$ (A) Mean probability to cooperate averaged over 100 independent lines of descent (average over 200 K updates, discarding the first 250 K and the last 50 K as described in section 3.2, as a function of synergy r for fixed (unevolvable) probability of punishment $p_P = 0.0, 0.25, 0.5, 0.75,$ and 1.0. The dashed line indicates a mean probability to cooperate $p_C = 0.5$, which we use to extrapolate the critical value r_c. This critical value depends on the punishment levels as predicted by equation (18). (B) Critical synergy r_c as a function of punishment probability p_P as deduced from panel (A) (points) by identifying the r_{crit} at which the cooperation probability $p_C = 0.5$. The dashed line indicates the prediction $r_c = 5(1 - \beta \rho_P)$, assuming that the density of cooperators $\rho_P \approx p_P$ (mean field), with $\beta = 0.8$ and $\gamma = 0.2$.
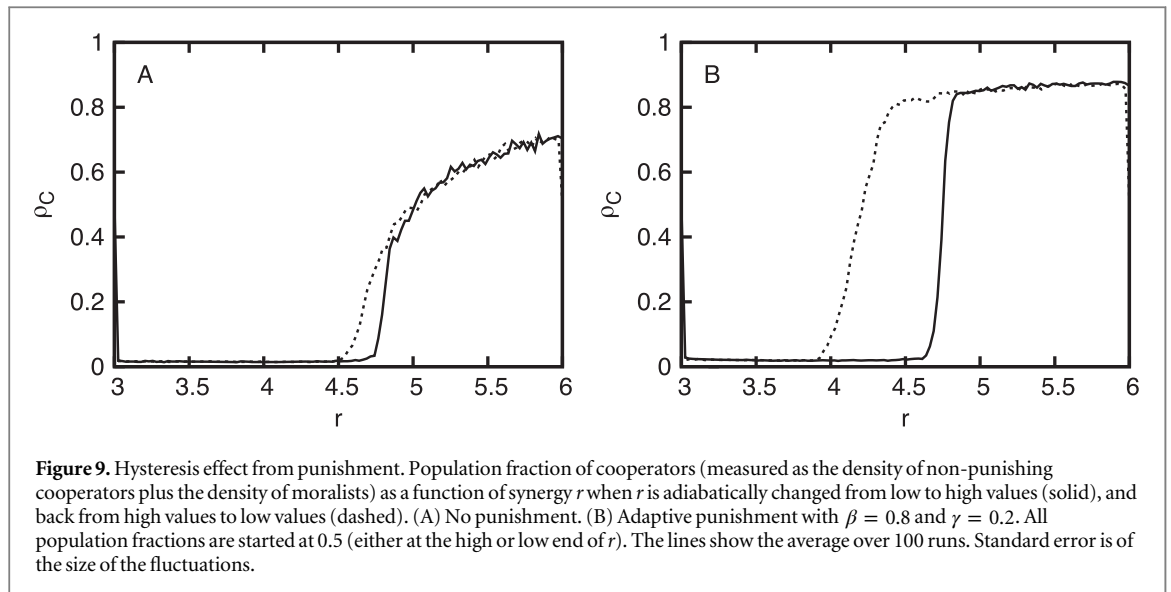
Because of the crucial importance of punishers in determining the synergy level at which cooperation emerges, the public goods game with a genetic basis (that is, with genes coding for probabilities of moves) implies curious dynamics close to the critical point. Below the critical point, defection is a stable strategy, and punishment is absent. When cooperation emerges as a possibility, punishment becomes more and more important, leading to a *lowering* of the critical synergy for cooperation via equation (18). At that point, cooperation emerges rapidly and decisively once a critical level of punishment has been achieved. Once cooperation is dominant and defectors are all but driven to extinction, punishment becomes irrelevant and the gene for punishment begins to drift. As this happens, the fraction of punishers drops, thus raising the critical synergy according to equation (18). As a consequence, a drifting punishment gene can lead to the sudden re-emergence of defectors as stable states. Once those have taken over, the reverse dynamics begins to unfold. Given this dynamic, we should observe periods of cooperation and defection that follow each other closely when the synergy is near the critical point.

These dynamics are reminiscent of the phenomenon of supercooling and superheating in certain phase transitions observed in condensed matter physics, as predicted in section 2. If we imagine the synergy parameter r as the critical parameter and the mean probability to cooperate as the order parameter, it is possible that when r is slowly increased, the population remains in the defecting phase because a switch to cooperation requires a critical number of cooperators as a 'seed'. In such a situation, the defecting phase is unstable to fluctuations. If a critical number of cooperators emerges by chance, punishment

immediately becomes effective against defectors, lowers the critical point as implied by equation (18), and the population could transition to cooperation very quickly. A hallmark of such bi-stable systems that require nucleation events in order to transition is *hysteresis*, a phenomenon where the state of the system depends on its history. We can test whether hysteresis exists in the public goods game (and whether the strength of this effect depends on the probability to punish), by adiabatically changing the synergy parameter first from low to high (transitioning from defection to cooperation), and then adiabatically back from high to low. While we see evidence of hysteresis even when punishment is absent (figure 9(A)), the effect is much more pronounced when punishment is possible (figure 9(B)). The population moves from cooperation to defection at about the expected critical synergy $r_{crit} \approx 4.15$ as r is decreased, but stays in the defecting phase much beyond the critical point as r is increased.

The observed hysteresis effect implies that once cooperation is established, it can be maintained even when the expected synergy fluctuates below the critical point, but that cooperation is difficult to establish even when the synergy would be conducive for that establishment. It also explains why levels of cooperation are higher when punishment is possible, even if punishment is used sparingly. In meta-stable phase transitions, bubbles of the new phase increase in size exponentially if larger than a critical size, but shrink exponentially when smaller than the critical size [52]. Thus, if a group invades with $\rho_C^{in} > \rho_{crit}$, $\rho_C \rightarrow 1$. This is different from the dynamics in the absence of punishment, where at the critical point all ρ_C^{in} have the same fitness, and the mean level of cooperation is 0.5, as is evident in figures 4 (inset), and 8(A). Indeed, the critical point for $\rho_P = 0$ (no punishment) is neutral,

Figure 9. Hysteresis effect from punishment. Population fraction of cooperators (measured as the density of non-punishing cooperators plus the density of moralists) as a function of synergy *r* when *r* is adiabatically changed from low to high values (solid), and back from high values to low values (dashed). (A) No punishment. (B) Adaptive punishment with $\beta = 0.8$ and $\gamma = 0.2$. All population fractions are started at 0.5 (either at the high or low end of *r*). The lines show the average over 100 runs. Standard error is of the size of the fluctuations.

while it is a repulsive fixed point when punishment is present. As a consequence, the phase transition as a function of *r* becomes steeper and steeper as punishment increases, and higher levels of cooperation are achieved. This behavior is strongly reminiscent of phase transitions in ferromagnetic systems where the 'rounding' of the transition due to impurities is reduced via the magnetic field. This suggests that a treatment of public goods games in terms of Ising-like models where punishment plays the role of a magnetic field forcing the alignment of spins should be possible, and we are currently pursuing such an approach [53].

5. Discussion

We studied the public goods game for well-mixed populations both theoretically and in agent-based simulations of Darwinian evolution of stochastic strategies, using genes that encode the probabilities for cooperation and punishment. It is known that punishment can drive the evolution of cooperation above a critical synergy level as long as there is a spatial structure in the environment [25, 26]. It was also previously believed that in well-mixed populations cooperation via punishment can only become successful if additional factors like reputation [22] or the potential for abstaining from the public good [29, 34] are influencing the evolution. Here we show that cooperation readily emerges in a well-mixed environment above a critical level of synergy. This critical level is influenced by a number of factors: the rate of punishment because punishment favors cooperating groups, but also spatial structure [25, 28, 29], because a single cooperator can nucleate a transition simply because offspring cooperators are placed next to it, giving rise to a 'bubble' of cooperators of sufficient size. This finding is similar to the observation of frequency-dependent cooperation when punishment (but not cooperation) is probabilistic [40], even

though the game studied by Chen *et al* is different from the game we consider here in that punishment is shared among the punisher, whereas in our game (which is the one studied by Helbing *et al* [26]), each punisher acts on his own.

We have not studied here the possibility of 'anti-social' punishment [19, 54, 55], where non-cooperating defectors can punish cooperators, but we do not expect this possibility to change the overall picture. Indeed, in simulations in which defection was not punished but instead rewarded (a negative punishment), this only served to reinforce the defecting phase, as was also found in [56]. A transition to the cooperative phase still takes place at sufficiently high synergy. Phase transitions between cooperative and defection phases have also been observed in a spatial version of the public goods game where costly rewards are given for cooperation, rather than the costly punishment for defectors [57]. It would be interesting to study that game within the context of the evolving probabilistic strategies studied here.

We conclude that in well-mixed populations cooperation can emerge if the synergy outweighs the defectors' reward, which is reduced by punishment. A punishment-dependent barrier to cooperation introduces an interesting dynamic near the critical synergy. Starting in the cooperative phase, as long as the mutation rate is low enough, the dearth of defectors in the cooperating phase makes punishment obsolete, that is, the selective pressure to punish disappears. As a consequence, the density of punishers decreases, thus increasing the critical point in turn. If the critical synergy has increased sufficiently, defectors can again gain a foothold. Such a shift, however, reinstates the selective pressure to punish, leading to a re-emergence of moralists that can drive defectors out once more. Thus, for synergy factors near the critical point, we can expect oscillations between cooperators and defectors, and no strategy is ever stable.

Finally, the observation of hysteresis implies the existence of metastable states, and gives rise to a self-enforcing (or 'self-aligning') dynamic where cooperation is stable even when punishment is never actually used. It is clear that meta-stable dynamics can only occur in the shaded region in figure 3, which is set by the probability π with which cooperators punish, and provides a mechanism to protect cooperating groups from defectors, as the defectors need to achieve a critical density in order to thrive. In a very real way, meta-stability raises the scepter of punishment to maintain cooperation, even when it is not used.

Acknowledgments

References

[1] Hardin G 1968 The tragedy of the commons *Science* **162** 1243–8

[2] Rankin D J, Bargum K and Kokko H 2007 The tragedy of the commons in evolutionary biology *Trends Ecol. Evol.* **22** 643–51

[3] MacLean R C 2008 The tragedy of the commons in microbial populations: insights from theoretical, comparative and experimental studies *Heredity* **100** 471–7

[4] Frank S 1996 Models of parasite virulence *Q. Rev. Biol.* **71** 37–78

[5] Brown S 1999 Cooperation and conflict in host-manipulating parasites *Proc. R. Soc.* B **266** 1899–904

[6] Frank S A 2006 *Foundations of Social Evolution* (Princeton, NJ: Princeton University Press)

[7] Maynard Smith J 1982 *Evolution and the Theory of Games* (Cambridge: Cambridge University Press)

[8] Axelrod R 1984 *The Evolution of Cooperation* (New York: Basic Books)

[9] Dugatkin L A 1997 *Cooperation Among Animals: An Evolutionary Perspective* (Princeton, NJ: Princeton University Press)

[10] Hofbauer J and Sigmund K 1998 *Evolutionary Games and Population Dynamics* (Cambridge: Cambridge University Press)

[11] Doebeli M and Hauert C 2005 Models of cooperation based on the prisoner's dilemma and the snowdrift game *Ecology Lett.* **8** 748–66

[12] Nowak M 2006 *Evolutionary Dynamics* (Cambridge, MA: Harvard University Press)

[13] Olson M 1971 *The Logic of Collective Action: Public Goods and the Theory of Groups* (Cambridge, MA: Harvard University Press)

[14] Davis D D and Holt C A 1993 *Experimental Economics* (Princeton, NJ: Princeton University Press)

[15] Ledyard J 1995 Public goods: a survey of experimental research *Handbook of Experimental Economics* ed J H Kagel and A E Roth (Princeton, NJ: Princeton University Press) pp 111–94

[16] Fehr E and Gächter S 2002 Altruistic punishment in humans *Nature* **415** 137–40

[17] Fehr E and Fischbacher U 2003 The nature of human altruism *Nature* **425** 785–91

[18] Hammerstein P (ed) 2003 *Genetic and Cultural Evolution of Cooperation* (Cambridge, MA: MIT Press)

[19] Nakamaru M and Iwasa Y 2006 The coevolution of altruism and punishment: role of the selfish punisher *J. Theor. Biol.* **240** 475–88

[20] Camerer C F and Fehr E 2006 When does 'economic man' dominate social behavior? *Science* **311** 47–52

[21] Gürerk O, Irlenbusch B and Rockenbach B 2006 The competitive advantage of sanctioning institutions *Science* **312** 108–11

[22] Sigmund K, Hauert C and Nowak M A 2001 Reward and punishment *Proc. Natl Acad. Sci. USA* **98** 10757–62

[23] Henrich J and Boyd R 2001 Why people punish defectors. Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas *J. Theor. Biol.* **208** 79–89

[24] Boyd R, Gintis H, Bowles S and Richerson P J 2003 The evolution of altruistic punishment *Proc. Natl Acad. Sci. USA* **100** 3531–5

[25] Brandt H, Hauert C and Sigmund K 2003 Punishment and reputation in spatial public goods games *Proc. Biol. Sci.* **270** 1099–104

[26] Helbing D, Szolnoki A, Perc M and Szabó G 2010 Evolutionary establishment of moral and double moral standards through spatial interactions *PLoS Comput. Biol.* **6** e1000758

[27] Boyd R, Gintis H and Bowles S 2010 Coordinated punishment of defectors sustains cooperation and can proliferate when rare *Science* **328** 617–20

[28] Helbing D, Szolnoki A, Perc M and Szabo G 2010 Punish, but not too hard: how costly punishment spreads in the spatial public goods game *New J. Phys.* **12** 083005

[29] Szabo G and Hauert C 2002 Phase transitions and volunteering in spatial public goods games *Phys. Rev. Lett.* **89** 118101

[30] Szolnoki A and Perc M 2013 Correlation of positive and negative reciprocity fails to confer an evolutionary advantage: phase transitions to elementary strategies *Phys. Rev.* X **3** 041021

[31] Perc M and Szolnoki A 2012 Self-organization of punishment in structured populations *New J. Phys.* **14** 043013

[32] Szolnoki A and Perc M 2013 Effectiveness of conditional punishment for the evolution of public cooperation *J. Theor. Biol.* **325** 34–41

[33] Fowler J H 2005 Altruistic punishment and the origin of cooperation *Proc. Natl Acad. Sci. USA* **102** 7047–9

[34] Hauert C, Traulsen A, Brandt H, Nowak M A and Sigmund K 2007 Via freedom to coercion: the emergence of costly punishment *Science* **316** 1905–7

[35] Silva H D, Hauert C, Traulsen A and Sigmund K 2009 Freedom, enforcement, and the social dilemma of strong altruism *J. Evol. Econ.* **20** 203–17

[36] Sigmund K, De Silva H, Traulsen A and Hauert C 2010 Social learning promotes institutions for governing the commons *Nature* **466** 861–3

[37] Yamagishi T 1986 The provision of a sanctioning system as a public good *J. Pers. Soc. Psychol.* **51** 110–6

[38] Sasaki T and Unemi T 2011 Replicator dynamics in public goods games with reward funds *J. Theor. Biol.* **287** 109–14

[39] Szolnoki A, Szabó G and Perc M 2011 Phase diagrams for the spatial public goods game with pool punishment *Phys. Rev.* E **83** 036101

[40] Chen X, Szolnoki A and Perc M 2014 Probabilistic sharing solves the problem of costly punishment *New J. Phys.* **16** 083016

[41] Pathria R 1996 *Statistical Mechanics* 2nd edn (Portsmouth, NH: Heinemann)

[42] Moran P A P 1962 *The Statistical Processes of Evolutionary Theory* (Oxford: Clarendon)

[43] Taylor C, Fudenberg D, Sasaki A and Nowak M A 2004 Evolutionary game dynamics in finite populations *Bull. Math. Biol.* **66** 1621–44

[44] Iliopoulos D, Hintze A and Adami C 2010 Critical dynamics in the evolution of stochastic strategies for the iterated Prisoner's Dilemma *PLoS Comput. Biol.* **7** e1000948

[45] Fletcher J A and Doebeli M 2009 A simple and general explanation for the evolution of altruism *Proc. Biol. Sci.* **276** 13–9

[46] Hintze A and Adami C 2010 Darwinian evolution of cooperation via punishment in the public goods game *Proc. 12th Int. Conf. of the Simulation and Synthesis of Living Systems* ed H Fellerman *et al* (Cambridge, MA: MIT Press) pp 445–550

[47] Lenski R E, Ofria C, Pennock R T and Adami C 2003 The evolutionary origin of complex features *Nature* **423** 139–44

[48] Østman B, Hintze A and Adami C 2012 Impact of epistasis and pleiotropy on evolutionary adaptation *Proc. R. Soc.* B **279** 247–56

[49] Imry Y and Wortis M 1979 Influence of quenched impurities on first-order phase transitions *Phys. Rev.* B 3580–5

[50] Zeeman E C 1981 Dynamics of the evolution of animal conflicts *J. Theor. Biol.* **89** 249–70

[51] Adami C, Schossau J and Hintze A 2012 Evolution and stability of altruist strategies in microbial games *Phys. Rev.* E **85** 011914

[52] Landau L D and Lifschitz E M 1987 *Statistische Physik Teil 1* (Berlin: Akademie-Verlag)

[53] Adami C, Pasmanter N and Hintze A 2015 Thermodynamics of evolutionary games (in preparation)

[54] Rand D G, Armao J J IV, Nakamaru M and Ohtsuki H 2010 Anti-social punishment can prevent the co-evolution of punishment and cooperation *J. Theor. Biol.* **265** 624–32

[55] Rand D G and Nowak M A 2011 The evolution of antisocial punishment in optional public goods games *Nat. Commun.* **2** 434

[56] Hauser O P, Nowak M A and Rand D G 2014 Punishment does not promote cooperation under exploration dynamics when anti-social punishment is possible *J. Theor. Biol.* **360** 163–71

[57] Szolnoki A and Perc M 2012 Evolutionary advantages of adaptive rewarding *New J. Phys.* **14** 093016