



# Synergy and Discounting of Cooperation in Social Dilemmas

## Citation

Hauert, Christoph, Franziska Michor, Martin A. Nowak, and Michael Doebeli. 2006. Synergy and discounting of cooperation in social dilemmas. *Journal of Theoretical Biology* 239(2): 195-202.

## Published Version

doi:10.1016/j.jtbi.2005.08.040

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:4317703>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# Synergy and discounting of cooperation in social dilemmas

Christoph Hauert<sup>\*,†,‡</sup>, Franziska Michor<sup>‡</sup>, Martin A. Nowak<sup>‡</sup> &  
Michael Doebeli<sup>†</sup>

<sup>†</sup> Departments of Zoology and Mathematics, University of British Columbia,  
6270 University Boulevard, Vancouver B.C. Canada, V6T 1Z4

<sup>‡</sup> Program for Evolutionary Dynamics,  
Department of Organismic and Evolutionary Biology,  
Department of Mathematics, Harvard University,  
One Brattle Square, Cambridge, MA 02138, USA

September 9, 2005

**Key words:** Evolution, Cooperation, Game Theory, Social Dilemmas

**Corresponding author:**

\* Christoph Hauert

Program for Evolutionary Dynamics, Harvard University

One Brattle Square, Cambridge, MA 02138, USA

email: [christoph\\_hauert@harvard.edu](mailto:christoph_hauert@harvard.edu)

phone: (617) 496 5550, fax: (617) 496 4629

**Manuscript information:** 23 pages, including references and 3 figures.

## Abstract

The emergence and maintainance of cooperation by natural selection is an enduring conundrum in evolutionary biology, which has been studied using a variety of game theoretical models inspired by different biological situations. The most widely studied games are the Prisoner's Dilemma, the Snowdrift game and by-product mutualism for pairwise interactions, as well as Public Goods games in larger groups of interacting individuals. Here we present a general framework for cooperation in social dilemmas in which all the traditional scenarios can be recovered as special cases. In social dilemmas cooperators provide a benefit to the group at some cost, while defectors exploit the group by reaping the benefits without bearing the costs of cooperation. Using the concepts of discounting and synergy for describing how benefits accumulate when more than one cooperator is present in a group of interacting individuals, we recover the four basic scenarios of evolutionary dynamics given by (i) dominating defection, (ii) co-existence of defectors and cooperators, (iii) dominating cooperation and (iv) bi-stability, in which cooperators and defectors cannot invade each other. Generically, for groups of three or more interacting individuals further, more complex, dynamics can occur. Our framework provides the first unifying approach to model cooperation in different kinds of social dilemmas.

# 1 Modeling Cooperation

The question of how natural selection can lead to cooperation has fascinated evolutionary biologists since Darwin (Darwin, 1859, Hammerstein, 2003, Maynard Smith & Szathmary, 1995, Trivers, 1971). Cooperation among relatives is explained by kin selection representing the idea that selfish genes lead to unselfish phenotypes (Frank, 1989, Hamilton, 1963). For the evolution of cooperation among genetically unrelated individuals various mechanisms have been put forward based on evolutionary game theory: cooperators form groups and thus preferentially interact with other cooperators (Sober & Wilson, 1998, Wilson & Sober, 1994); cooperators occupy spatial positions in lattices or networks and interact with their neighbors who are also cooperators (Hauert, 2001, Killingback *et al.*, 1999, Nowak & May, 1992); optional interactions can stabilize cooperation (Hauert *et al.*, 2002, Semmann *et al.*, 2003); repeated games enable the emergence of direct reciprocity (Axelrod & Hamilton, 1981, Nowak & Sigmund, 1993); and reputation facilitates the evolution of cooperation via indirect reciprocity (Alexander, 1987, Nowak & Sigmund, 1998) or punishment opportunities (Sigmund *et al.*, 2001).

The vast majority of models on the evolution of cooperation consider pairwise interactions: a cooperator meeting another cooperator obtains the reward  $R$ , but against a defector the cooperator is left with the sucker’s payoff  $S$ . In contrast, the defector exploits the cooperator and receives the temptation  $T$ , but when facing another defector each gets the punishment  $P$ . This terminology was introduced for the Prisoner’s Dilemma, which is defined by the payoff ranking  $T > R > P > S$ . Hence defection dominates cooperation because it is better to defect regardless of what the partner does. In terms of costs  $c$  and benefits  $b$  of cooperation, the Prisoner’s Dilemma describes situations where costs incur to the cooperator but benefits accrue exclusively to the partner, i.e.  $T = b$ ,  $R = b - c$ ,  $P = 0$  and  $S = -c$  with  $b > c$ . This represents the most stringent form of a social dilemma.

## 1.1 Social Dilemmas

Social dilemmas are defined as interactions in groups of individuals where groups of cooperators are better off than groups of defectors, but in any mixed group defectors outperform cooperators (Dawes, 1980). The fact that defectors exploit cooperators requires that defectors are better off in any mixed group than in the absence of cooperators, and conversely that cooperators are worse off in any mixed group than in the absence of defectors. In the context of pairwise interactions this requires  $R > P, T > S, T > P$  and  $R > S$ . These requirements are satisfied by the Prisoner's Dilemma but there are three additional rankings possible.

The Snowdrift game (Sugden, 1986) (also known as the Chicken or Hawk-Dove game (Maynard Smith & Price, 1973)) is defined by  $T > R > S > P$ . To illustrate this game, consider two drivers on their way home that are caught in a blizzard and trapped on either side of a snowdrift. Each driver has the option to remove the snowdrift and start shoveling or to remain in the car. In contrast to the Prisoner's Dilemma, the best choice now clearly depends on the other driver: if the other cooperates and starts shoveling, it pays to defect and remain in the car but if the other defects, it is better to shovel and get home than to wait for spring. Similar situations may occur whenever the act of cooperation creates a common good that can be exploited by others, i.e. whenever the benefits of cooperation accrue not only to the partner but also to the cooperator itself. For example, foraging yeast cells produce and secrete an enzyme in order to lyse their environment (Greig & Travisano, 2004). The resulting food resource represents a valuable common good prone to exploitation by other cells. However, despite the prospects of being exploited, a single cell may be better off (prevent starvation) by producing the enzyme if no one else does. This last twist relaxes the social dilemma to some extent.

The dilemma is completely relaxed if  $R > T > S > P$  holds. This refers to by-product mutualism where it is better to cooperate irrespective of the partner's

decision, i.e. cooperation dominates defection. However, note that in any mixed group, defectors are still better off than cooperators but at the same time the payoff of defecting individuals would increase upon switching to cooperation (because the individual draws a net benefit from its own cooperative act). The remaining ranking is  $R > T > P > S$ , for which the best choice again depends on the partner's decision, but now it is best to aim for mutual decisions: defect if the other defects and cooperate if the other cooperates. The social dilemma presents itself as a coordination problem with mutual cooperation as the preferred outcome.

Cooperative interactions in groups of  $N$  individuals have received much attention in the context of Public Goods games (Fehr & Gächter, 2002, Kagel & Roth, 1995). In typical Public Goods experiments, individuals can make an investment into a common pool knowing that the experimenter will multiply the total investments and distribute it equally among all participants irrespective of their contributions. In essence, Public Goods games represent  $N$ -persons Prisoner's Dilemmas (Dugatkin, 1997, Hauert & Szabó, 2003), and defection invariably dominates cooperation. As in pairwise interactions, maintenance of cooperation requires additional mechanisms such as iterated interactions (Boyd & Richerson, 1988, Hauert & Schuster, 1997) or local interactions in spatially structured populations (Hauert & Doebeli, 2004, Pollock, 1989). Group interactions under less stringent conditions have hardly been studied.

## 1.2 Evolutionary Dynamics in Groups of $N$ Individuals

In evolutionary dynamics we consider infinitely large populations consisting of a fraction of cooperators  $x$  and  $1 - x$  defectors. According to replicator dynamics (Hofbauer & Sigmund, 1998), changes in  $x$  are determined by the relative performance of cooperators as compared to defectors:

$$\dot{x} = x(1 - x)(f_C - f_D) \tag{1}$$

where  $f_C, f_D$  denote the average fitness, i.e. the average payoff, of cooperators and defectors, respectively. Here we present a general framework to study interactions of cooperators and defectors in groups of  $N$  players. Cooperators are defined as individuals that provide a benefit to all members of the group at some cost to themselves. In contrast, defectors attempt to exploit the common enterprise, avoid the costs and provide no benefit. We denote by  $P_C(k)$  and  $P_D(k)$ , respectively, the payoff for a single cooperator and a single defector in a group that contains  $k$  cooperators. Note that for the payoff of a cooperator,  $P_C(k)$ , the cooperator is one of the  $k$  cooperators in the group. Thus,  $P_C(k)$  is defined for  $k = 1, \dots, N$ , while  $P_D(k)$  is meaningful for  $k = 0, \dots, N - 1$ . Also note that, in accordance with social dilemmas,  $P_C(k) < P_D(k)$  must hold for all  $k = 1, \dots, N - 1$ , so that in any given group each defector has a higher payoff than each cooperator.

In groups that are formed at random according to binomial sampling, the average fitness of cooperators and defectors is given by

$$f_C = \sum_{j=0}^{N-1} \binom{N-1}{j} x^j (1-x)^{N-1-j} P_C(j+1) \quad (2a)$$

$$f_D = \sum_{j=0}^{N-1} \binom{N-1}{j} x^j (1-x)^{N-1-j} P_D(j). \quad (2b)$$

Here  $\binom{N-1}{j} x^j (1-x)^{N-1-j}$  is the probability that there are  $j$  cooperators among the  $N - 1$  other individuals in a group of size  $N$  in which the focal cooperator or defector finds itself. Consequently, the formulas above represent weighted averages of the payoffs of a focal cooperator (focal defector) facing  $j$  cooperators among the  $N - 1$  co-players.

## 2 A Framework of Cooperation

In natural systems, the actual value of the benefits provided by cooperators may depend on the number of cooperators in the group. For example, in the case of

foraging yeast cells, the benefit provided by the first cooperator may be critical for survival, whereas the value of additional food decreases until, eventually, more food becomes useless because the cells are saturated. Similarly, cooperators may produce enzymes for enzyme-mediated reactions. The efficiency of the reaction may exhibit a faster than linear increase with concentration such that additional enzyme production has an enhanced value (Fersht, 1977, Hammes, 1982). This leads to discounted or synergistically enhanced benefits based on the number of cooperators in groups of interacting individuals.

## 2.1 Synergy and Discounting

In order to model the synergistically enhanced as well as the discounted net value of accumulated cooperative benefits, we assume that defectors and cooperators, respectively, receive the following payoffs in a group with  $k$  cooperators:

$$P_D(k) = \frac{b}{N}(1 + w + w^2 + \dots + w^{k-1}) = \frac{b}{N} \frac{1 - w^k}{1 - w} \quad (3a)$$

$$P_C(k) = P_D(k) - c. \quad (3b)$$

Hence, the first cooperator provides a benefit  $b$  which is shared by all  $N$  members of the group (including itself), the second one increases everyone's benefit by  $b/N \cdot w$ , and so on, to the last of the  $k$  cooperators in the group providing an additional benefit of  $b/N \cdot w^{k-1}$ . The costs of cooperation  $c$ , however, incur only to cooperators. If  $w = 1$ , then all cooperators provide the same incremental benefit  $b/N$ . This corresponds to the traditional formulation of Public Goods games with  $P_D(k) = rkc/N$  (or  $b = rc$ ) where  $r$  denotes the multiplication factor of the common pool. If  $w < 1$ , then the benefits are discounted and the value of the benefits provided by each additional cooperator is lower than the previous one. If  $w > 1$ , then the benefits are synergistically enhanced, and each additional cooperator provides incremental benefits of increasing value. At this point it is important to note that neither discounting



nor synergy involve temporal components, i.e. this does not refer to potential future benefits in iterated interactions, which is an entirely different line of research with strong roots in economics (Fudenberg & Maskin, 1986).

Substituting the payoff functions Eq. (3) into Eq. (2) determines the average performance of cooperators and defectors in a population with a fraction  $x$  of cooperators:

$$f_C = \frac{b}{N(1-w)} \left(1 - w(1-x+wx)^{N-1}\right) - c \quad (4a)$$

$$f_D = \frac{b}{N(1-w)} \left(1 - (1-x+wx)^{N-1}\right). \quad (4b)$$

It is easy to see from these expressions that the equation  $f_C(x^*) = f_D(x^*)$  has at most one solution  $x^*$  in the open interval  $(0, 1)$ , which, if it exists, is given by  $x^* = [1 - (cN/b)^{1/(N-1)}]/[1 - w]$  (see Fig. 1). Therefore, the replicator dynamics (see Eq. (1)) has at most one interior equilibrium. It follows that the dynamics can be determined by considering the invasion capabilities of cooperators and defectors, i.e. the stability of the two trivial equilibria  $x = 0$  and  $x = 1$ .

## 2.2 Classification of Social Dilemmas

Based on these calculations, we obtain a natural classification of the dynamics in  $N$ -player group interactions, which generates the four basic scenarios of evolutionary dynamics (Nowak & Sigmund, 2004):

- (i) Defectors dominate cooperators if  $cN/b > 1$  and  $cN/b > w^{N-1}$  because  $f_D > f_C$  holds (Fig. 1a). The only stable equilibrium is  $x = 0$ . Note that for  $cN/b > 1$  the minimal benefit secured by a cooperator (arising from its own act of cooperation) does not exceed the incurring costs. This parameter region corresponds to the Prisoner's Dilemma or Public Goods games.
- (ii) If  $1 > cN/b > w^{N-1}$  the two equilibria  $x = 0$  and  $x = 1$  are both unstable. Cooperators and defectors can invade each other and coexist at a stable equi-

librium  $x^*$ . This is reflected by  $f_C > f_D$  for  $x < x^*$  but  $f_C < f_D$  for  $x > x^*$  (Fig. 1b). This parameter region represents a generalization of the Snowdrift game to groups of  $N$  players.

(iii) Cooperators dominate defectors if  $cN/b < 1$  and  $cN/b < w^{N-1}$  because  $f_C > f_D$  holds (Fig. 1c). Note, however, that defectors are still better off in every (mixed) group ( $P_D(k) > P_C(k)$ ), but each defector could further increase its payoff by switching to cooperation ( $P_C(k+1) > P_D(k)$ ). No interior equilibrium exists and the only stable equilibrium is  $x = 1$ . In this parameter region the social dilemma is completely relaxed and cooperation evolves through by-product mutualism.

(iv) If  $w^{N-1} > cN/b > 1$  the two equilibria  $x = 0$  and  $x = 1$  are both stable, i.e. cooperators and defectors cannot invade each other. The basins of attraction of the two equilibria are separated by the unstable equilibrium  $x^*$ . In this bistable situation the evolutionary outcome depends on the initial fraction of cooperators  $x_0$ : if  $x_0 > x^*$  cooperation evolves ( $f_C > f_D$ ) but if  $x_0 < x^*$  cooperation vanishes ( $f_D > f_C$ ) (Fig. 1d).

Note that the conditions for successful invasions can be directly obtained from  $P_C(k)$  and  $P_D(k)$ , independent of binomial (or any other) sampling. The reason for this is that if cooperators are rare, then they can only increase in abundance if a single cooperator in a group of  $N - 1$  defectors has a higher payoff than a group of  $N$  defectors. Thus, cooperators can invade if  $P_C(1) > P_D(0)$ . Likewise, if defectors are rare, they invade only if  $P_D(N - 1) > P_C(N)$ , i.e. if a single defector in a group with  $N - 1$  cooperators receives a higher payoff than a cooperator in a group of  $N$  cooperators. These conditions are sufficient to generate the above classification. Scenarios (i) and (ii), i.e. dominant defection and co-existence, are reminiscent of the distinction between 'strong' and 'weak' altruism (Charlesworth, 1979, Wilson, 1979)

where the act of cooperation incurs either absolute costs to the cooperator ( $b/N < c$ ) or only relative costs as compared to the performance of defectors ( $b/N > c$ ).

The different dynamical regimes are summarized in Fig. 2 using phase diagrams. The figure highlights the fact that transitions between the qualitatively different types of evolutionary games can be generated by varying the parameters  $w$ ,  $c/b$ , and  $N$ . Not surprisingly, synergy ( $w > 1$ ) generally favours cooperation as compared to discounting ( $w < 1$ ), whereas increasing the cost-to-benefit ratio  $c/b$  or the group size  $N$  generally favours defection (Fig. 2). For discounted benefits,  $w < 1$ , defection reigns for  $b/N < c$  whereas for  $b/N > c$  cooperators persist. In the latter case, the critical group size given by  $n_c = 1 + \log(cN/b)/\log(w)$  separates by-product mutualism and the (generalized) Snowdrift game: for  $N < n_c$  cooperators dominate defectors, whereas for  $N > n_c$  cooperators and defectors coexist (Fig. 2b). For synergistic interactions,  $w > 1$ , cooperation dominates for  $b/N > c$  whereas for  $b/N < c$  defectors dominate if  $N > n_c$  but if  $N < n_c$  the dynamics becomes bistable (Fig. 2b).

The suggested discount/synergy framework can be fully analyzed regardless of the group size  $N$  because there exists at most a single interior equilibrium  $x^*$ . Since this already holds for  $N = 2$  no qualitatively different scenarios are found for larger groups of interacting individuals (even though group size does have an effect on the dynamics, see Fig. 2). The analysis hinges on the fact that the value of the benefits provided by  $k$  cooperators in a group of  $N$  interacting players are captured in the single parameter  $w$ .

### 2.3 Extending the Framework

More generally, one could, for example, introduce  $N$  different parameters  $\alpha_j$ ,  $j = 1, \dots, N$  describing these effects, so that  $P_D(k) = \alpha_1 + \dots + \alpha_{k-1}$  (and e.g. again  $P_C(k) = P_D(k) - c$ ). It is easy to see that, in principle, such systems can have

anywhere between zero and  $N - 1$  interior equilibria because  $f_C, f_D$  are polynomials in  $x$  of degree  $N - 1$ . Moreover, there are up to  $2N$  different dynamical scenarios:  $N$  different arrangements of interior equilibria with two alternative configurations each because the stability of adjacent equilibria must alternate.

### Competition among Defectors

To illustrate the possibility of multiple interior equilibria in a biologically motivated setting, we first consider a situation that includes interactions among defectors. Let us assume a vicious type of defector that not only avoids the costs of cooperation but additionally competes with other defectors for their share of the benefit. The strength of competition increases with the number of defectors in a group. For simplicity we neglect competition among cooperators because if all individuals compete, the overall benefit of cooperation would essentially decrease. Defector competition can be described by changing their payoff to

$$P_D(k) = \frac{b}{N} \frac{1 - w^k}{1 - w} u^{N-1-k} \quad (5)$$

i.e.

$$f_D = \frac{b}{N(1-w)} \left( (u - ux + x)^{N-1} - (u - ux + wx)^{N-1} \right)$$

where  $0 < u < 1$  measures how the strength of competition increases with the number of defectors: the defector payoff decreases due to the presence of  $N - 1 - k$  other defectors in the group, and this decrease is more pronounced for smaller  $u$ . The cooperator payoff is assumed to be unchanged (see Eq. (3)). With this payoff structure, analytical solutions are no longer attainable in general, but Fig. 3 demonstrates that the existence of a second interior equilibrium point can lead to dynamic regimes not seen in the case without defector interaction. This is illustrated in Fig. 3a,b for  $N = 3$  and  $b/N < c$ . Competition among defectors supports cooperators such that

they survive over a broader range of parameters. Moreover, this allows for stable co-existence of cooperators and defectors for  $w < 1$  and  $b/N < c$ , which is otherwise impossible (c.f. Fig. 2). Note that for  $b/N > c$  the qualitative dynamics with  $u < 1$  is the same as with  $u = 1$ , i.e., as in absence of defector interactions.

### Variations in Effects of Discounting and Synergy

Multiple interior equilibria also occur if cooperators and defectors use different discounting/synergy factors  $w$  and  $v$ , i.e. the value of the common good provided by the cooperators is different for the two types. This could occur for example if the two strategy types correspond to different physiological states such that cooperators and defectors differ in their efficiency in taking advantage of the common resource. We assume that the cooperator payoff remains the same as before (see Eqs. (3) and (4)), and that the defector payoff becomes

$$P_D(k) = \frac{b}{N} \frac{1 - v^k}{1 - v}, \quad (6)$$

i.e.

$$f_D = \frac{b}{N(1 - v)} \left( 1 - (1 - x + vx)^{N-1} \right).$$

Only  $v = w$  allows for an analytical solution of  $f_C = f_D$  but in general there are again up to  $N - 1$  equilibria in  $(0, 1)$ . This case is illustrated in Fig. 3c,d for  $N = 3$ . For example, for  $b/N > c$  large  $v$  increase the domain of co-existence, whereas small  $v$  promote dominance of cooperation. More complicated scenarios, including coexistence of multiple stable interior equilibria, can be observed for larger  $N$ .

## 3 Conclusions

In this paper, we present a general framework for cooperation in evolutionary  $N$ -player games that encompasses and recovers traditional games such as the Prisoner's

Dilemma or Public Goods games as special cases. The basis of our framework is formed by the concept of discounting and synergy, which simply takes into account that the actual value of the benefits provided by cooperators may depend on the total number of cooperators in the group. Thus, with discounting the value of the benefit provided by the first cooperator in a group is  $b$ , but the value of the benefits provided by each additional cooperator is discounted by a factor  $w < 1$  as compared to the previous cooperator. All cooperators pay a cost  $c$ . For example, this leads to the definition of an  $N$ -player Snowdrift game if the discounting factor  $w$  and the cost-benefit ratio  $cN/b$  are sufficiently small (Fig. 2). It is important to recall that discounting does not refer to potential future benefits but rather to the process of accumulating benefits provided by multiple cooperators. In the case of synergy, the value of the benefit provided by each additional cooperator is synergistically enhanced by a factor  $w > 1$ .

Viewing the traditional games from the perspective of this general framework emphasizes that the various scenarios - Prisoner's Dilemma or Public Goods games, Snowdrift games, by-product mutualism, and bistability - are interconnected through variations of the continuous parameters  $w$ ,  $c/b$  and the group size  $N$ , which seamlessly relates seemingly disparate biological situations. For example, the discomfort with the Prisoner's Dilemma as the sole model for cooperation is increasing (Clutton-Brock, 2002, Heinsohn & Parker, 1995, West *et al.*, 2002), but viewing cooperation in the framework of discounting opens up natural connections to related scenarios, such as the Snowdrift game. In this way, our framework could prove to be helpful in bridging the gap between theoretical advances and experimental evidence.

In experimental settings it is notoriously difficult to quantitatively assess the fitness of strategic/behavioral patterns. For example, sticklebacks inspect their predators preferably in pairs and are believed to be trapped in a Prisoner's Dilemma (Milinski, 1987). However, despite tremendous efforts, only the payoff ranking  $T > R > S$

has been experimentally confirmed (Milinski *et al.*, 1997). Consequently, it remains unresolved whether the fish indeed engage in a Prisoner's Dilemma (requiring  $R > P > S$ ) or rather in a Snowdrift game ( $R > S > P$ ). In another example, a Prisoner's Dilemma interaction has been shown to occur between RNA phages within host cells (Turner & Chao, 1999), but selection alters the payoff structure such that cooperative and defective phage strains coexist in a Snowdrift game (Turner & Chao, 2003).

Similarly, it has been argued that Prisoner's Dilemma interactions occur in the aforementioned case of enzyme production in foraging yeast cells (Greig & Travisano, 2004). Despite the apparent connection to the Prisoner's Dilemma game given by the possibility of cheating, such frequency dependent benefits may be better captured by the Snowdrift game: if cooperators abound, defection is dominant and selfish individuals exploit the accrued benefits but as cooperators become rare, the costly enzyme production, may provide sufficient advantage to the producing individual despite by-product benefits to others, such that cooperation becomes dominant. Indeed, Greig & Travisano (2004) report that cheating was beneficial only if a substantial fraction of the yeast population was cooperating, i.e., producing the enzyme.

Outside of biology, the study of social dilemmas has received particular attention by experimental economists and anthropologists (Fehr & Gächter, 2002, Henrich *et al.*, 2001, Panchanathan & Boyd, 2004). Humans display an apparently irrational, high readiness to cooperate in Public Goods and Ultimatum games (Güth *et al.*, 1982, Nowak *et al.*, 2000), which confounds the basic rationality assumptions of *homo oeconomicus*. In both games, defection is dominant but the Ultimatum game adds aspects of punishment because it can be interpreted as a Prisoner's Dilemma interaction followed by a round of (costly) punishment (Sigmund *et al.*, 2001). Punishment and reputation have been identified as very potent promoters of human cooperation in social dilemmas (Fehr & Gächter, 2002, Milinski *et al.*, 2002, Wedekind & Milinski,

2000). Such additional mechanisms can be easily incorporated into our framework. However, already in Public Goods interactions, where only the multiplication factor of the common good depends on the total amount invested, qualitatively different outcomes can be generated, which allow e.g. for co-existence of cooperators and defectors in a generalized Snowdrift game. In experimental settings, variations of the multiplication factor could test the sensitivity of human behavior to quantitative and qualitative changes of the interaction characteristics.

In summary, Snowdrift games can be considered as social dilemmas that are intermediate between Prisoner's Dilemma games (or Public Goods games in larger groups) and by-product mutualism, which occur whenever ordinary selfish behavior benefits others (Brown, 1983, West-Eberhard, 1975). By-product mutualism has also been put forth to challenge the Prisoner's Dilemma for explaining patterns of cooperation in natural populations (Connor, 1995, 1996, Dugatkin, 1996, Milinski, 1996). Our general theoretical framework for cooperation in social dilemmas seems capable of reconciling the different viewpoints and emphasizes that the different dynamical domains of social dilemmas are related by continuous changes in biologically meaningful parameters.



## References

- Alexander, R. D. (1987). *The Biology of Moral Systems*. Aldine de Gruyter, New York.
- Axelrod, R. & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, **211**, 1390–1396.
- Boyd, R. & Richerson, P. J. (1988). The evolution of reciprocity in sizeable groups. *J. theor. Biol.* **132**, 337–356.
- Brown, J. L. (1983). Cooperation: a biologist's dilemma. *Adv. Study Behav.* **13**, 1–37.
- Charlesworth, B. (1979). A note on the evolution of altruism in structured demes. *Am. Nat.* **113** (4), 601–605.
- Clutton-Brock, T. (2002). Breeding together: kin selection and mutualism in cooperative vertebrates. *Science*, **296**, 69–72.
- Connor, R. C. (1995). Altruism among non-relatives: alternatives to the 'Prisoner's Dilemma'. *TREE*, **10** (2), 84–86.
- Connor, R. C. (1996). Partner preferences in by-product mutualisms and the case of predator inspection in fish. *Anim. Behav.* **51**, 451–454.
- Darwin, C. (1859). *The Origin of Species*. Harvard University Press, 1964, Reprinted.
- Dawes, R. M. (1980). Social dilemmas. *Ann. Rev. Psychol.* **31**, 169–193.
- Dugatkin, L. A. (1996). Tit for tat, by-product mutualism and predator inspection: a reply to connor. *Anim. Behav.* **51**, 455–457.

- Dugatkin, L. A. (1997). *Cooperation among animals: an evolutionary perspective*. Oxford University Press, Oxford.
- Fehr, E. & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, **415**, 137–140.
- Fersht, A. (1977). *Enzyme Structure and Mechanism*. W. H. Freeman & Co, San Francisco.
- Frank, R. H. (1989). *Passions Within Reason*. W. W. Norton & Company.
- Fudenberg, D. & Maskin, E. (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, **54** (3), 533–554.
- Greig, D. & Travisano, M. (2004). The Prisoner's Dilemma and polymorphism in yeast SUC genes. *Biol. Lett.* **271**, S25 – S26.
- Güth, W., Schmittberger, R. & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* **3**, 376–388.
- Hamilton, W. D. (1963). The evolution of altruistic behaviour. *Am. Nat.* **97**, 354–356.
- Hammerstein, P., ed. (2003). *Genetic and Cultural Evolution of Cooperation*. MIT Press, Cambridge Mass.
- Hammes, G. G. (1982). *Enzyme Catalysis and Regulation*. Academic Press, New York.
- Hauert, C. (2001). Fundamental clusters in spatial  $2 \times 2$  games. *Proc. R. Soc. Lond. B*, **268**, 761–769.
- Hauert, C., De Monte, S., Hofbauer, J. & Sigmund, K. (2002). Volunteering as red queen mechanism for cooperation in public goods games. *Science*, **296**, 1129–1132.

- Hauert, C. & Doebeli, M. (2004). Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, **428**, 643–646.
- Hauert, C. & Schuster, H. G. (1997). Effects of increasing the number of players and memory size in the iterated prisoner's dilemma: a numerical approach. *Proc. R. Soc. Lond. B*, **264**, 513–519.
- Hauert, C. & Szabó, G. (2003). Prisoner's dilemma and public goods games in different geometries: compulsory versus voluntary interactions. *Complexity*, **8** (4), 31–38.
- Heinsohn, R. & Parker, C. (1995). Complex cooperative strategies in group-territorial african lions. *Science*, **269**, 1260–1262.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. & McElreath, R. (2001). Cooperation, reciprocity and punishment in fifteen small scale societies. *Am. Econ. Rev.* **91**, 73–78.
- Hofbauer, J. & Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge.
- Kagel, J. H. & Roth, A. E., eds (1995). *The handbook of experimental economics*. Princeton University Press, Princeton.
- Killingback, T., Doebeli, M. & Knowlton, N. (1999). Variable investment, the continuous prisoner's dilemma, and the origin of cooperation. *Proc. R. Soc. Lond. B*, **266**, 1723–1728.
- Maynard Smith, J. & Price, G. (1973). The logic of animal conflict. *Nature*, **246**, 15–18.
- Maynard Smith, J. & Szathmáry, E. (1995). *The major transitions in evolution*. W. H. Freeman & Co, Oxford.

- Milinski, M. (1987). Tit for tat in sticklebacks and the evolution of cooperation. *Nature*, **325**, 433–435.
- Milinski, M. (1996). By-product mutualism, Tit-for-Tat and cooperative predator inspection: a reply to Connor. *Anim. Behav.* **51**, 458–461.
- Milinski, M., Lüthi, J. H., Eggler, R. & Parker, G. A. (1997). Cooperation under predation risk: experiments on costs and benefits. *Proc. R. Soc. Lond. B*, **264**, 831–837.
- Milinski, M., Semmann, D. & Krambeck, H.-J. (2002). Reputation helps solve the 'tragedy of the commons'. *Nature*, **415**, 424–426.
- Nowak, M. A. & May, R. M. (1992). Evolutionary games and spatial chaos. *Nature*, **359**, 826–829.
- Nowak, M. A., Page, K. M. & Sigmund, K. (2000). Fairness versus reason in the ultimatum game. *Science*, **289**, 1773–1775.
- Nowak, M. A. & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, **364**, 56–58.
- Nowak, M. A. & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, **393**, 573–577.
- Nowak, M. A. & Sigmund, K. (2004). Evolutionary dynamics of biological games. *Science*, **303**, 793–799.
- Panchanathan, K. & Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature*, **432**, 499–502.
- Pollock, G. B. (1989). Evolutionary stability of reciprocity in a viscous lattice. *Social Networks*, **11**, 175–212.

- Semmann, D., Krambeck, H.-J. & Milinski, M. (2003). Volunteering leads to rock-paper-scissors dynamics in a public goods game. *Nature*, **425**, 390–393.
- Sigmund, K., Hauert, C. & Nowak, M. A. (2001). Reward and punishment. *Proc. Natl. Acad. Sci.* **98**, 10757–10762.
- Sober, E. & Wilson, D. S. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Harvard University Press, Cambridge, MA.
- Sugden, R. (1986). *The Economics of Rights, Co-operation and Welfare*. Blackwell, Oxford and New York.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57.
- Turner, P. E. & Chao, L. (1999). Prisoner's dilemma in an RNA virus. *Nature*, **398**, 441–443.
- Turner, P. E. & Chao, L. (2003). Escape from prisoner's dilemma in RNA phage  $\phi 6$ . *Am. Nat.* **161**, 497–505.
- Wedekind, C. & Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, **288**, 850–852.
- West, S. A., Pen, I. & Griffin, A. S. (2002). Cooperation and competition between relatives. *Science*, **296**, 72–75.
- West-Eberhard, M. J. (1975). The Evolution of Social Behavior by Kin Selection. *Q. Rev. Biol.* **50**, 1–33.
- Wilson, D. S. (1979). Structured demes and trait-group variation. *Am. Nat.* **113** (4), 606–610.
- Wilson, D. S. & Sober, E. (1994). Reintroducing group selection to the human behavioral sciences. *Behav. Brain Sci.* **17**, 585–654.

# Figures

hauert\_figure\_1

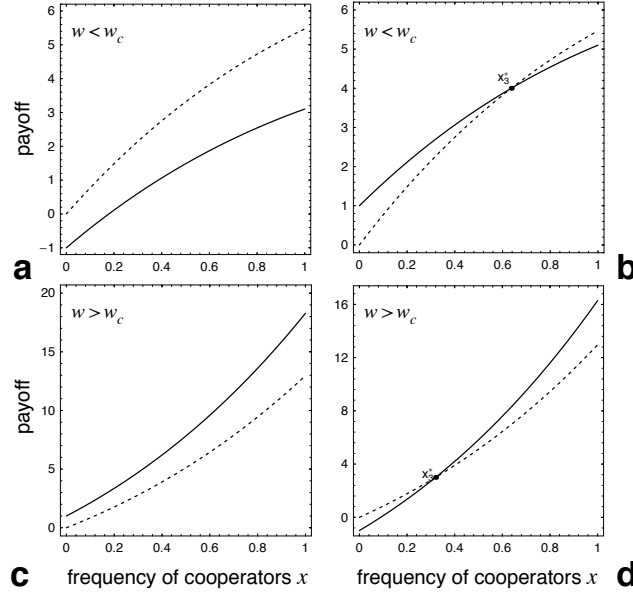


Figure 1: Full classification of the evolutionary dynamics for discounted and synergistically enhanced benefits of cooperation in social dilemma interactions in groups of arbitrary size. The average payoff of cooperators,  $f_C$  (solid line), and defectors,  $f_D$  (dashed line) is shown as a function of the frequency of cooperators  $x$  for different discount/synergy factors  $w$ . **a** Defection dominates for  $b/N < c$  and  $w < w_c = (c/b)^{1/(N-1)}$  ( $f_D > f_C$  for all  $x$ ). This corresponds to Public Goods (Prisoner's Dilemma) interactions. **b** For larger benefits ( $b/N > c, w < w_c$ ) the game dynamics has a stable equilibrium  $x^*$  ( $f_C > f_D$  for  $x < x^*$  but  $f_C < f_D$  for  $x > x^*$ ). This generalizes the Snowdrift game to interactions in groups of arbitrary size. **c** For larger  $w > w_c$  ( $b/N > c$ ) cooperation becomes dominant ( $f_C > f_D$  for all  $x$ ). This scenario describes by-product mutualism. **d** For  $b/N < c$  and  $w > w_c$  the dynamics is bi-stable: for  $x < x^*$  cooperators vanish ( $f_D > f_C$ ) but for  $x > x^*$  cooperators dominate ( $f_C > f_D$ ). Parameters:  $N = 5, b = 10$  and **a**  $c = 3, w = 3/4$ ; **b**  $c = 1, w = 3/4$ ; **c**  $c = 1, w = 4/3$ ; **d**  $c = 3, w = 4/3$ .

## hauert\_figure\_2

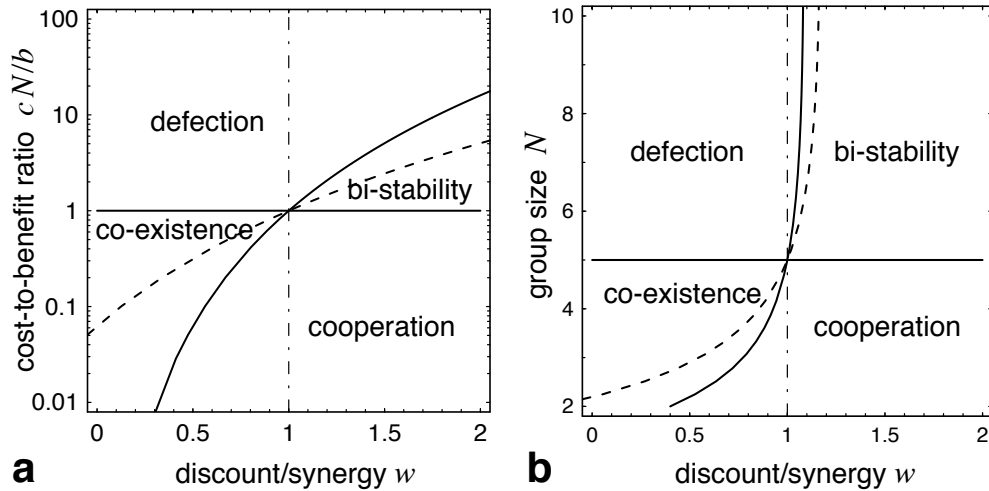


Figure 2: Phase diagrams illustrating the different dynamical regimes. The dash-dotted line separates discounted and synergistically enhanced benefits ( $w = 1$ ). **a** For any fixed group size (here  $N = 5$ ) the dynamics is determined by the cost-to-benefit ratio  $c/b$  and the discount/synergy factor  $w$ . In the case of discounting ( $w < 1$ ) decreasing  $c/b$  facilitates cooperation and may lead to transitions from dominant defection (Prisoner's Dilemma) to co-existence of cooperators and defectors (Snow-drift games) and finally to dominant cooperation (by-product mutualism). Along the dashed line the interior fixed point is  $x^* = 0.5$ . Similarly, in the case of synergy ( $w > 1$ ), increasing the benefits turns dominant defection into bi-stability and ends with dominant cooperation. Note that for  $w < 1$  cooperators can only survive if  $cN/b < 1$ . This condition is not necessary for  $w > 1$ . Also note that coexistence requires  $w < 1$  and bi-stability  $w > 1$ , whereas dominant cooperation or defection can occur for  $w$  greater, less or equal to one. Increasing  $N$  leaves the qualitative results unchanged and merely increases the slope of the boundary between co-existence and cooperation (defection and bi-stability) but it always runs through  $cN/b = w = 1$ . **b** Dynamics as determined by group size  $N$  and discount/synergy  $w$  for  $c/b = 2/10$ . Decreasing  $N$  favors cooperation by increasing the domain of cooperation, decreasing the domain of defection and, additionally, by shifting the co-existence equilibrium towards cooperation as well as by increasing the basin of attraction of the cooperative state in the case of bi-stable dynamics. The dashed line again marks  $x^* = 0.5$ .

hauert\_figure\_3

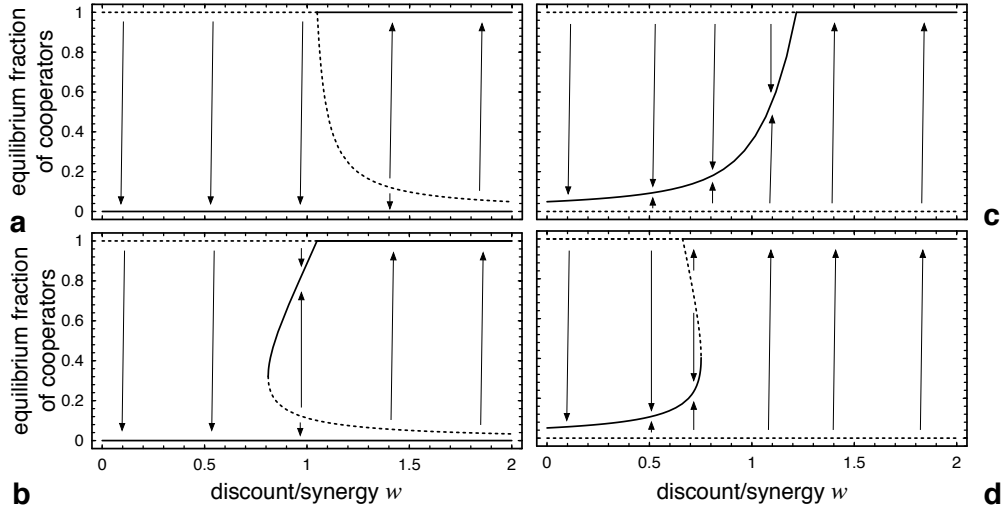


Figure 3: Multiple interior equilibria can occur in groups of  $N \geq 3$  interacting individuals. Positions of stable (solid line) and unstable (dashed line) equilibria are shown as a function of the discount/synergy factor  $w$ , which acts as a bifurcation parameter. **a** An unstable interior fixed point appears upon increasing  $w$  and the dynamics changes from dominating defection to bi-stability ( $b/N < c$ ). **b** Same as **a** but including competition among defectors within each group (see Eq. (5)). This introduces a saddle-node bifurcation as  $w$  is increased, such that cooperators either go extinct or co-exist with a small fraction of defectors. For further increases in  $w$  the stable interior equilibrium leaves the interval  $(0, 1)$  resulting in a transcritical bifurcation which leads to bi-stability. **c** Cooperators and defectors have different discount/synergy factors  $w$  and  $v$ , respectively (see Eq. (6)). For  $b/N > c$  and larger  $v$  the region of co-existence extends into the realm of synergy ( $w > 1$ ). This never happens in the case  $v = w$  (c.f. Fig. 2). **d** Same as **c** but for smaller  $v$ : another saddle-node bifurcation occurs as  $w$  is decreased, which changes the regime of dominant cooperation into a regime of bi-stability, where pure cooperation as well as co-existence with a minority of cooperators are stable. Decreasing  $w$  further leads to another transcritical bifurcation as the unstable equilibrium leaves  $(0, 1)$  and results in globally stable coexistence. Parameters:  $N = 3, b = 3$  and **a**  $c = 1.1$ ; **b**  $c = 1.1, u = 0.5$ ; **c**  $c = 0.9, v = 1.8$ ; **d**  $c = 0.9, v = 0.2$ .