Supporting Text:
Exploration Dynamics in Evolutionary Games

Arne Traulsen, Christoph Hauert, Hannelore De Silva (née Brandt),
Martin A. Nowak, and Karl Sigmund

January 5, 2009

1 Payoffs

First, we calculate the average payoffs for the public goods game with all four strategic types. From this, it is straightforward to infer the payoffs if some of the strategies are absent. In one particular interaction with i_C cooperators, i_D defectors, i_L loners, and i_P punishers, we have $S = i_C + i_P + i_D$ participants. For $S \leq 1$ no interaction takes place and everybody receives the loners payoff σ. For $S > 1$, the four different types obtain the payoffs

$$\begin{aligned}
P_C &= r\,c\frac{i_C + i_P}{S} - c, \\
P_D &= r\,c\frac{i_C + i_P}{S} - \beta i_P, \\
P_L &= \sigma, \\
P_P &= r\,c\frac{i_C + i_P}{S} - c - \gamma i_D.
\end{aligned} \tag{1}$$

The average payoffs π_k of the four strategies can be calculated from the average group composition. In a population with X_C cooperators, X_D defectors, X_L loners, and X_P punishers ($X_C+X_D+X_L+X_P = M$), a randomly sampled group of size N has a particular composition with probability given by the multivariate hypergeometric distribution,

$$H(\boldsymbol{i}, \boldsymbol{X}) = \frac{\binom{X_C}{i_C}\binom{X_D}{i_D}\binom{X_L}{i_L}\binom{X_P}{i_P}}{\binom{M}{N}}. \tag{2}$$

Here, $\boldsymbol{i} = (i_C, i_D, i_L, i_P)$ and $\boldsymbol{X} = (X_C, X_D, X_L, X_P)$. The average group composition determines the average payoffs $\pi_k = \sum_{i_C}\sum_{i_D}\sum_{i_L}\sum_{i_P} H(\ldots)P_k$, which simplify to

$$\pi_C = B(\boldsymbol{X}) - F(X_L)c \tag{3}$$

$$\pi_D = B(\boldsymbol{X}) - \frac{X_P}{M-1}(N-1)\beta \tag{4}$$

$$\pi_L = \sigma \tag{5}$$

$$\pi_P = B(\boldsymbol{X}) - F(X_L)c - \frac{X_D}{M-1}(N-1)\gamma. \tag{6}$$

The quantity $B(\boldsymbol{X})$ (the return of the public goods game or σ in the case of only one participant) is given by

$$\begin{aligned}
B(\boldsymbol{X}) \quad = \quad & \frac{rc(X_C + X_P)}{M - X_L - 1}\left[1 - \frac{M}{N(M - X_L)}\right] \\
+ \quad & \frac{\binom{X_L}{N-1}}{\binom{M-1}{N-1}}\left[\sigma + \frac{rc(X_C + X_P)(X_L - N + 1)}{N(M - X_L - 1)(M - X_L)}\right].
\end{aligned} \tag{7}$$

The effective cost to contribute to the public good is

$$\begin{aligned}
F(X_L) \quad = \quad & 1 - \frac{r}{N}\frac{M - N}{M - X_L - 1} \\
+ \quad & \frac{\binom{X_L}{N-1}}{\binom{M-1}{N-1}}\left(\frac{r}{N}\frac{X_L + 1}{M - X_L - 1} + r\frac{M - X_L - 2}{M - X_L - 1} - 1\right).
\end{aligned} \tag{8}$$

The relevant payoffs for the replicator equation are obtained in the limit $M \to \infty$ or from an equivalent calculation assuming infinite M from the beginning [1].

As mentioned in the main text, to avoid that punishment trivially succeeds (if D is dominated by P) or trivially fails (if P is dominated by D), the pairwise comparison of the punisher and defector strategy requires bistability. In other words, a single punisher cannot invade a population of defectors and, conversely, a single defector cannot invade a population of punishers. This is satisfied if a single punisher has a smaller payoff than the resident defectors:

$$c\left(\frac{r}{N}-1\right)-\gamma(N-1)<\frac{N-1}{M-1}\left(\frac{rc}{N}-\beta\right). \tag{9}$$

For $\beta < (M-1)\gamma$ this always holds. This is a weak condition and gets violated only if punishment becomes very cheap. Conversely, a single defector has a smaller payoff than the resident punishers if

$$(N-1)\beta > \frac{N-1}{M-1}\gamma + c\left(1-\frac{r}{N}\frac{M-N}{M-1}\right). \tag{10}$$

For large populations ($M \gg N$), this reduces to $(N-1)\beta > c(1-\frac{r}{N})$. In particular, if the punishment exceeds the maximal costs of cooperation this always holds.

2 Evolutionary dynamics

Here, the two different analytical approximations for the evolutionary dynamics in the limits of small and high mutation or exploration rates are discussed in detail.

2.1 Small mutation rates

In finite populations with small mutation rates [2, 3], the population is homogeneous most of the time. Occasionally a mutation occurs with probability μ and an individual switches to a different random strategy. The mutant either reaches fixation or extinction before the next mutant arises. The average time until a neutral mutant reaches fixation is $M(M-1)$ time steps [4]. The average time between two mutations is μ^{-1}. Thus, for $\mu \ll M^{-2}$, the time scales of mutation and imitation are separated and we only have to consider two strategies at a time. The probability of fixation can be calculated analytical for any birth death process [5], but becomes particularly simple in our case of strong selection: 0, 1, or $\frac{1}{M}$ (for neutral transitions).

If only cooperators and defectors are present, the probability that a single cooperator takes over a defector population is zero for strong selection. A single defector, however, always takes over a cooperator population. This leads to the following transition matrix between the

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 1-\mu & \mu \\ 0 & 1 \end{pmatrix} \end{array}. \tag{11}$$

The first line determines the probability that the population remains in the C state or ends up in the D state after a mutation. With probability $1-\mu$, no mutation occurs and the population remains in C. With probability μ, a mutation leads to a D individual, which takes over and brings the system to the D state. Mutations in the D state can produce C individuals, but they cannot take over. Therefore, the probability to move from D to C is zero. The stationary distribution is given by the eigenvector corresponding to the eigenvalue 1, which is $(x_C, x_D) = (0, 1)$. Thus, the system stays in the D state.

For public goods games with punishment, we obtain the transition matrix

$$\begin{array}{cc} & \begin{array}{ccc} C & D & P \end{array} \\ \begin{array}{c} C \\ D \\ P \end{array} & \begin{pmatrix} 1-\frac{\mu}{2}-\frac{\mu}{M} & \frac{\mu}{2} & \frac{\mu}{2M} \\ 0 & 1 & 0 \\ \frac{\mu}{2M} & 0 & 1-\frac{\mu}{2M} \end{pmatrix} \end{array}. \tag{12}$$

Let us discuss the first line again, i.e. the population is in the C state: If a mutation occurs, the mutant is a punisher with probability $\frac{1}{2}$ and takes over the population with probability $\frac{1}{M}$. This leads to the entry in the last line, $\frac{\mu}{2M}$. With probability $\frac{1}{2}$, the mutant is a defector, which always takes over the population. Thus, the probability to go from the cooperator state to the defector state is $\frac{\mu}{2}$. The probability to go from C to C follows from normalization of the first line. The transition probabilities in the remaining lines follow equivalently. The eigenvector corresponding to the eigenvalue 1 is $(x_C, x_D, x_P) = (0, 1, 0)$. Thus, the system again always remains in the defector state.

For voluntary public goods games, i.e. with loners instead of punishers, the transition matrix reads

$$\begin{array}{cc} & \begin{array}{ccc} C & D & L \end{array} \\ \begin{array}{c} C \\ D \\ L \end{array} & \begin{pmatrix} 1-\frac{\mu}{2} & \frac{\mu}{2} & 0 \\ 0 & 1-\frac{\mu}{2} & \frac{\mu}{2} \\ \frac{\mu}{4} & 0 & 1-\frac{\mu}{4} \end{pmatrix} \end{array}. \tag{13}$$

Let us discuss the first entry in the last line. The probability to go from the loner state to the cooperator state is $\frac{\mu}{4}$, because such a mutation occurs with probability $\frac{\mu}{2}$ and the mutant takes over with probability $\frac{1}{2}$. The remaining entries follow analogous reasoning as above. The stationary distribution yields $(x_C, x_D, x_L) = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$.

The transition matrix for all four strategies is

$$\begin{array}{cc} & \begin{array}{cccc} C & D & L & P \end{array} \\ \begin{array}{c} C \\ D \\ L \\ P \end{array} & \begin{pmatrix} 1-\frac{\mu}{3}-\frac{\mu}{3M} & \frac{\mu}{3} & 0 & \frac{\mu}{3M} \\ 0 & 1-\frac{\mu}{3} & \frac{\mu}{3} & 0 \\ \frac{\mu}{6} & 0 & 1-\frac{\mu}{3} & \frac{\mu}{6} \\ \frac{\mu}{3M} & 0 & 0 & 1-\frac{\mu}{3M} \end{pmatrix} \end{array}. \tag{14}$$

We obtain the stationary distribution from the eigenvector corresponding to the eigenvalue 1 of the stochastic transition matrix, $(x_C, x_D, x_L, x_P) = \frac{1}{8+M}(2, 2, 2, 2+M)$. For large populations, punishers prevail, $(x_C, x_D, x_L, x_P) \to (0, 0, 0, 1)$. Even though the system spends a negligible amount of time in the loner state, they are pivotal in order to tip the scale in favor of cooperation (and punishment) [2]. This follows directly from the comparison with the C, D, P case.

3 High mutation rates

We state the procedure for the general case of d strategies here. From the payoffs π of the strategies, we derive the transition probabilities for the imitation process. The probability to choose an individual of type j and transform it into type k either through imitation or mutation is

$$T_{j\to k}(\boldsymbol{X}) = \frac{X_j}{M}\left((1-\mu)\frac{X_k}{M}\Theta\left[\pi_k - \pi_j\right] + \frac{\mu}{d-1}\right). \tag{15}$$

Here, $j, k = C, D, L, P$ and $\Theta[x]$ denotes the Heavyside function. The equation can be rationalized as follows: With probability $\frac{X_j}{M}$, an individual of type j is selected as focal individual. With probability $1-\mu$, no mutation occurs. The role model is of type k with probability $\frac{X_k}{M}$. Switching to the role model's strategy occurs only for $\pi_k > \pi_j$. With probability μ, a mutation occurs and the focal individual switches to one of the $d-1$ alternative strategies with equal probability. Thus, the probability to adopt strategy k is $\frac{1}{d-1}$.

The Master equation governing the dynamics of the probability $\Omega^t(\boldsymbol{X})$ that the system is in state $\boldsymbol{X}$ is given by

$$\Omega^{t+1}(\boldsymbol{X}) = \Omega^t(\boldsymbol{X}) - T^{\rm out}(\boldsymbol{X})\Omega^t(\boldsymbol{X}) + T^{\rm in}(\boldsymbol{X}+\boldsymbol{\delta})\Omega^t(\boldsymbol{X}+\boldsymbol{\delta}), \tag{16}$$

where $T^{\rm out}(\boldsymbol{X}) = \sum_{j,k,j\neq k} T_{j\to k}(\boldsymbol{X})$ summarizes the transition probabilities leading away from $\boldsymbol{X}$. Similarly, the transitions from neighboring states $\boldsymbol{X}+\boldsymbol{\delta}$ into $\boldsymbol{X}$ are symbolized by $T^{\rm in}(\boldsymbol{X}+\boldsymbol{\delta})$. Equivalently to $T^{\rm out}(\boldsymbol{X})$, also $T^{\rm in}(\boldsymbol{X})$ involves a sum, but here the sum is over the neighboring states, which makes the notation more complex. Details can be found in [6]. Employing a Kramers-Moyal expansion of the Master equation for large system size M, we can derive a Fokker-Planck equation for the probability density $\rho(\boldsymbol{x})$ [6, 7, 8],

$$\dot{\rho}(\boldsymbol{x}) = -\sum_k \frac{\partial}{\partial x_k}\rho(\boldsymbol{x})\mathcal{A}_k(\boldsymbol{x}) + \frac{1}{2}\sum_{j,k}\frac{\partial^2}{\partial x_k \partial x_j}\rho(\boldsymbol{x})\mathcal{B}_{jk}(\boldsymbol{x}) \tag{17}$$

where $x_k = X_k/M$. The drift vector describing the deterministic part of the dynamics is given by

$$\mathcal{A}_k(\boldsymbol{x}) = \sum_j (T_{j\to k} - T_{k\to j}) \tag{18}$$

and the diffusion matrix characterizing the noise by

$$\mathcal{B}_{jk}(\boldsymbol{x}) = \frac{1}{M}[-(T_{j\to k} + T_{k\to j}) + \delta_{jk}\sum_l (T_{j\to l} + T_{l\to j})]. \tag{19}$$

The noise scales with M^{-1} and determines the variances of the distribution for large M. Note that the game has no impact on the diffusion term because $T_{j\to l} + T_{l\to j} = (1-\mu)x_j x_l + \mu(x_j + x_l)/(d-1)$. Moreover, mutations affect the properties of the noise, but $\mathcal{B}_{jk}(\boldsymbol{x})$ neither vanishes for $\mu = 0$ nor for $\mu = 1$.

The stationary probability distribution has maxima close to the points with $\mathcal{A}_k(\boldsymbol{x}) = 0$. These points are computed numerically. Several such points may appear when decreasing μ, but the noise typically selects one of them. For numerical procedures, we approximate the Heavyside-function $\Theta[x]$ by the Fermi function $F(x) = \left[1 + e^{-x/T}\right]^{-1}$, where $T \ll 1$.

In the case of two strategies, e.g. cooperators and defectors, the drift term is particularly simple and reduces to $\mathcal{A}_C(\boldsymbol{x}) = T_{D\to C}(\boldsymbol{x}) - T_{C\to D}(\boldsymbol{x})$. The solution of $\mathcal{A}_C(\boldsymbol{x}) = 0$ reads in this special case

$$x_C = \frac{1 + \mu - \sqrt{1 - 2\mu + 5\mu^2}}{2 - 2\mu}. \tag{20}$$

For the general case with $d > 2$, $\mathcal{A}_C(\boldsymbol{x}) = 0$ can only be solved numerically.

References

[1] Brandt, H, Hauert, C, Sigmund, K (2006) Punishing and abstaining for public goods. *Proc. Natl. Acad. Sci. USA* 103:495–497.

[2] Hauert, C, Traulsen, A, Brandt, H, Nowak, MA, Sigmund, K (2007) Via freedom to coercion: the emergence of costly punishment. *Science* 316:1905–1907.

[3] Imhof, LA, Fudenberg, D, Nowak, MA (2005) Evolutionary cycles of cooperation and defection. *Proc. Natl. Acad. Sci. USA* 102:10797–10800.

[4] Antal, T, Scheuring, I (2006) Fixation of strategies for an evolutionary game in finite populations. *Bull. Math. Biol.* 68:1923–1944.

[5] Nowak, MA (2006) *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA).

[6] Traulsen, A, Claussen, JC, Hauert, C (2006) Coevolutionary dynamics in large, but finite populations. *Phys. Rev. E* 74:11901.

[7] Helbing, D (1993) Boltzmann-like and Boltzmann-Fokker-Planck equations as a foundation of behavioral models. *Physica A* 196:546–573.

[8] Traulsen, A, Claussen, JC, Hauert, C (2005) Coevolutionary dynamics: From finite to infinite populations. *Phys. Rev. Lett.* 95:238701.