

Stochastic evolutionary dynamics of direct reciprocity

Lorens A. Imhof^{1,*} and Martin A. Nowak²

¹*Department of Economics, Hausdorff Center for Mathematics, Bonn University, 53113 Bonn, Germany*

²*Program for Evolutionary Dynamics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA*

Evolutionary game theory is the study of frequency-dependent selection. The success of an individual depends on the frequencies of strategies that are used in the population. We propose a new model for studying evolutionary dynamics in games with a continuous strategy space. The population size is finite. All members of the population use the same strategy. A mutant strategy is chosen from some distribution over the strategy space. The fixation probability of the mutant strategy in the resident population is calculated. The new mutant takes over the population with this probability. In this case, the mutant becomes the new resident. Otherwise, the existing resident remains. Then, another mutant is generated. These dynamics lead to a stationary distribution over the entire strategy space. Our new approach generalizes classical adaptive dynamics in three ways: (i) the population size is finite; (ii) mutants can be drawn non-locally and (iii) the dynamics are stochastic. We explore reactive strategies in the repeated Prisoner's Dilemma. We perform 'knock-out experiments' to study how various strategies affect the evolution of cooperation. We find that 'tit-for-tat' is a weak catalyst for the emergence of cooperation, while 'always cooperate' is a strong catalyst for the emergence of defection. Our analysis leads to a new understanding of the optimal level of forgiveness that is needed for the evolution of cooperation under direct reciprocity.

Keywords: cooperation; finite population; game theory; mathematical biology; Prisoner's Dilemma

1. INTRODUCTION

The basic concept in evolutionary game dynamics is that the payoff from a game leads to reproductive success (Maynard Smith & Price 1973; Maynard Smith 1982; Hofbauer & Sigmund 1998; Cressman 2003; Nowak & Sigmund 2004). Reproduction can be genetic or cultural. If there is a small number of discrete strategies, then the replicator dynamics provide a powerful framework for studying deterministic dynamics in infinitely large populations (Taylor & Jonker 1978; Weibull 1995; Samuelson 1997; Fudenberg & Levine 1998; Hofbauer & Sigmund 2003). There is a formal equivalence between the replicator equation and the Lotka–Volterra equation of ecology (Hofbauer & Sigmund 1998), thereby providing an important bridge between evolutionary game theory and mathematical ecology (May 1973). This equivalence is not surprising, because the success of species in an ecosystem depends on the abundance of other species. The replicator equation is one of the several dynamical frameworks that can be used, and other update rules have been proposed (Ritzberger & Weibull 1995; Fudenberg & Levine 1998; Hofbauer & Sigmund 1998). Recently, there is also much effort to explore the stochastic evolutionary game dynamics in populations of finite size (Schaffer 1988; Kandori *et al.* 1993; Nowak *et al.* 2004; Imhof *et al.* 2005; Tao & Cressman 2007; Santos *et al.* 2008).

If the strategy space is continuous, then there is no simple and perfect way to study evolutionary dynamics.

A widely used approach is given by adaptive dynamics (Hofbauer & Sigmund 1990; Nowak & Sigmund 1990; Dieckmann & Law 1996; Metz *et al.* 1996; Dieckmann *et al.* 2000), which rests on the following assumptions: there is an infinitely large population; at any one time, the individuals use one (or a few) strategies; mutants are drawn from an infinitesimally small region around the resident strategy; invasion and fixation dynamics are deterministic. Adaptive dynamics represent a useful tool for studying games with continuous strategy space and have led to many important insights (Doebeli *et al.* 2004; Ackermann *et al.* 2008).

2. MODEL

Here, we propose a new way to study evolutionary dynamics for games over a continuous strategy space. In our approach, there is a finite population. All the individuals in this population use the same strategy. Then, a mutant strategy occurs that can be drawn from an arbitrary distribution over the strategy space. This distribution can be constant or it can depend on the resident strategy. Next, we consider the stochastic game dynamics between resident and mutants. We calculate the fixation probability, ρ , that the mutant will take over the population. As a particular model for the stochastic dynamics, we consider a frequency-dependent Moran process (Nowak *et al.* 2004), but other models have been investigated and could also be used here (Binmore & Samuelson 1997; Traulsen *et al.* 2005; Vukov & Szabó 2005; Imhof & Nowak 2006; Lessard & Ladret 2007). With probability ρ , the resident is replaced by the mutant. With probability $1 - \rho$, the invasion is repelled

* Author for correspondence (limhof@uni-bonn.de).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2009.1171> or via <http://rspb.royalsocietypublishing.org>.

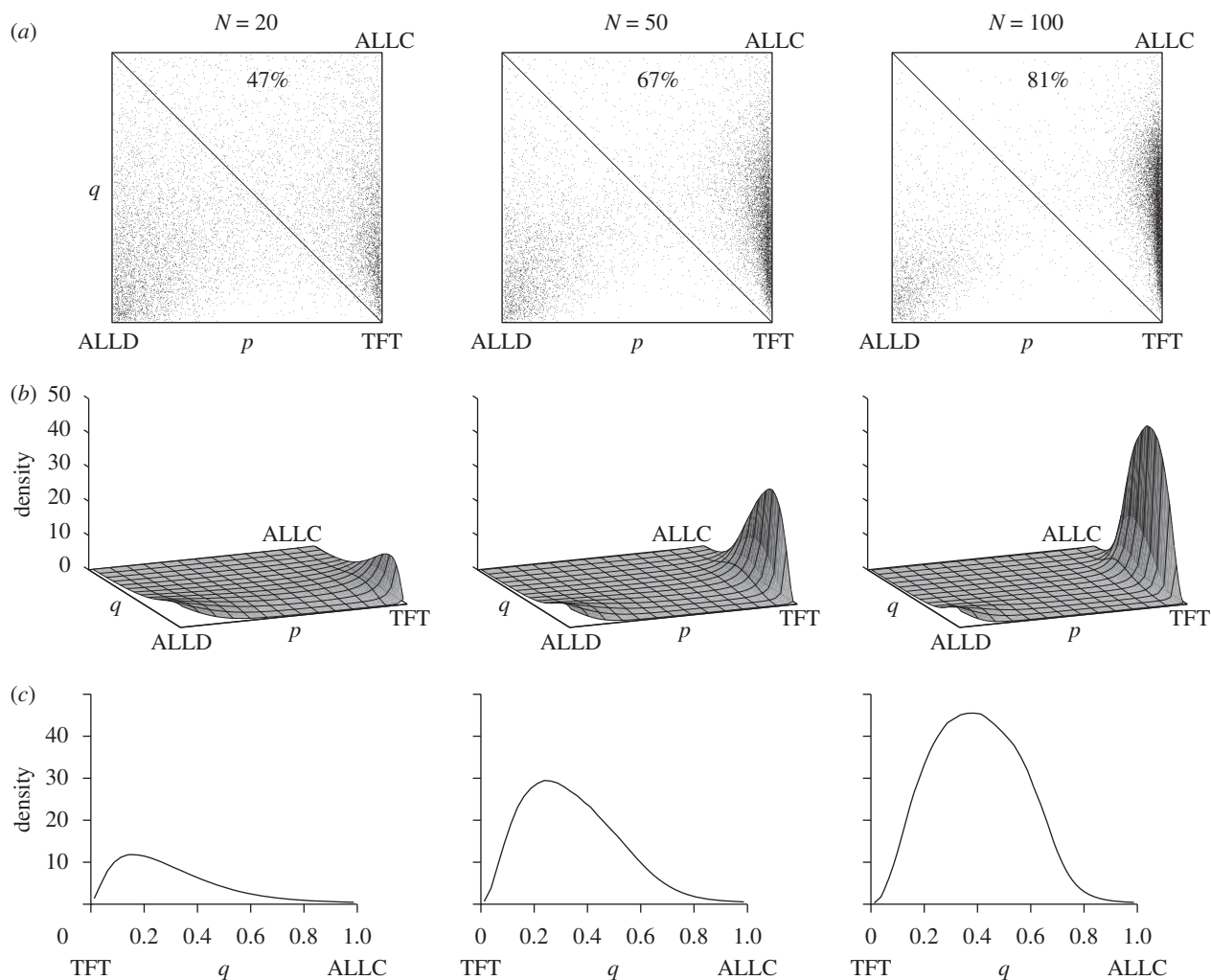


Figure 1. Stationary distributions of stochastic evolutionary dynamics. The state space is the set of reactive strategies (p, q) in the iterated Prisoner's Dilemma, where p is the conditional probability of cooperating after the opponent has cooperated in the previous round and q is the conditional probability of cooperating after the opponent has defected in the previous round. The stationary distribution describes the proportion of time that the population spends in a given part of the state space. New mutants are chosen from a uniform distribution on the state space. Parameter values are $b = 10$ and $c = 1$. (a) Sample points drawn from the stationary distributions for selected population sizes N . Darker parts indicate the areas visited more frequently. The per cent numbers indicate the proportion of time that the population is in the cooperative half, where $p + q > 1$. (b) Densities of the stationary distributions. The concentration of the distribution near ALLD and GTFT becomes larger as the population size increases. (c) Sections of the stationary densities along the TFT-ALLC edge. This edge describes the set of all GTFT strategies where $p = 1$ and the parameter q specifies the degree of forgiveness. For our parameters, the classical value is $q = 1 - c/b = 0.9$, a strategy which occurs only rarely in our finite population model. The q values observed most often depend on the population size and increase from about 0.15 for $N = 20$ to about 0.37 for $N = 100$.

and the resident remains. Then, another mutant is generated and the algorithm enters the next cycle.

If we repeat these steps many times, then the evolutionary dynamics will lead to a stationary distribution over the space of all strategies. This stationary distribution is informative about many aspects of the evolutionary dynamics of the game that is being considered. The maxima of this distribution signify successful strategies. Success is a combination of two factors: (i) it depends on how able a strategy is in resisting invasion attempts and (ii) it depends on how able a strategy is in invading other (successful) strategies. Both properties depend on the pairwise fixation probabilities. Proulx & Day (2001) studied a similar finite population model to demonstrate the relevance of fixation probabilities in stochastic environments.

As a fascinating example, we study the evolution of cooperation under direct reciprocity (Trivers 1971;

Axelrod & Hamilton 1981; Fudenberg & Maskin 1990; Nowak & Sigmund 1990, 1992). We investigate the simplified Prisoner's Dilemma, where cooperation means paying a cost, c , for the other individual to receive a benefit, b ; defection implies no cost and provides no benefit. Direct reciprocity can arise, if there are repeated encounters between the same two individuals. My strategy depends on what has happened in the previous encounters. In particular, my strategy depends on what you have done to me.

We analyse the set of reactive strategies (Nowak & Sigmund 1990), which are given by two parameters: p is the probability of cooperating, if the other person has cooperated in the previous round; q is the probability of cooperating, if the other person has defected in the previous round. Each reactive strategy is given by a pair (p, q) . This set contains some well-known strategies: for example, always defect, ALLD, $(0,0)$; always cooperate, ALLC, $(1,1)$; and tit-for-tat, TFT, $(1,0)$.

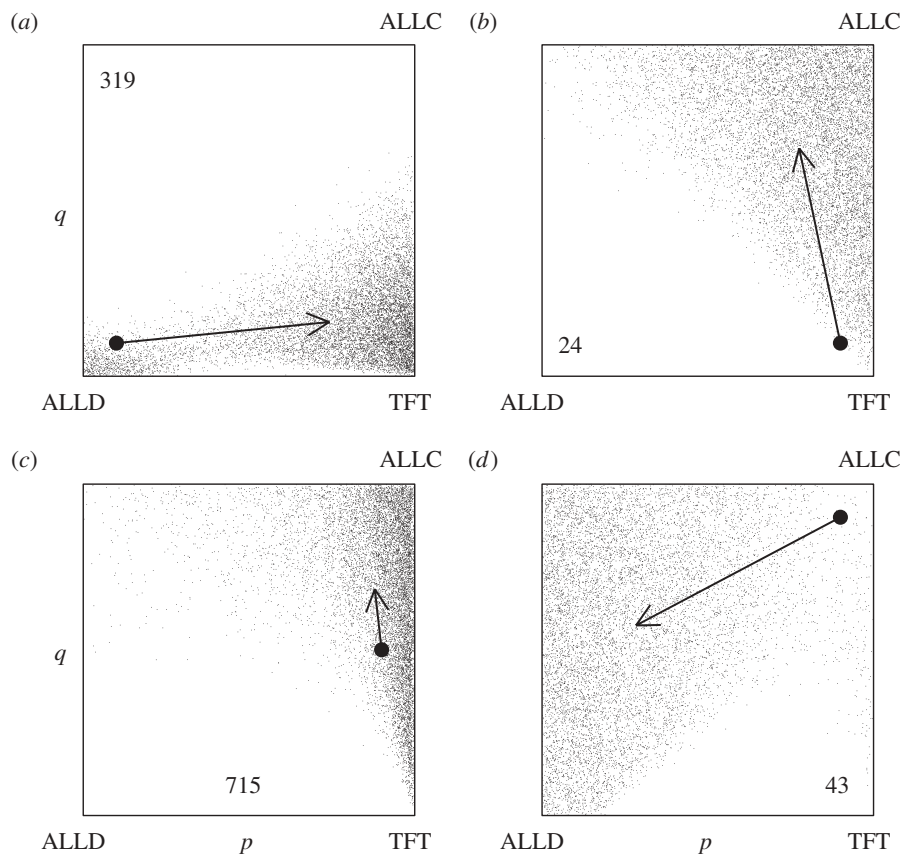


Figure 2. Successful invaders. The big black point is the resident strategy. The cloud shows the distribution of the first successful mutation that replaces the resident. The arrowhead indicates the centre of gravity of the cloud of successful invaders. The number quantifies how many unsuccessful mutants are generated on average before one successful mutant takes over. (a) Typically, an ALLD population is taken over after a while by a mutant with TFT-like behaviour. (b) A TFT population is then quickly taken over and the dynamics move towards some new state close to the edge from TFT to ALLC. (c) If the new resident strategy is carefully forgiving—in this example we have $(p, q) = (0.9, 0.5)$ —the population stays at this state for a long time, but will eventually move closer to ALLC. (d) An ALLC population is quickly taken over and the population returns to ALLD. As in figure 1, mutations are generated from a uniform distribution. Parameter values are population size $N = 100$, benefit $b = 10$ and cost $c = 1$.

Previous explorations of this strategy set (Nowak & Sigmund 1992), which were based on the replicator equation, have revealed the following evolutionary dynamics. From a random ensemble of strategies, ALLD emerges victorious. Subsequently, TFT catalyses the transition to cooperative societies. But TFT enjoys only a brief moment of glory, because it cannot correct mistakes: in the presence of noise, the expected payoff of two TFT players is low. TFT is soon replaced by generous tit-for-tat, GTFT, which always cooperates when the other person has cooperated, but sometimes even cooperates when the other person has defected. Therefore, GTFT can correct mistakes: it is given by $(1, q)$, where the magnitude of the parameter q is usually defined as the maximum value of forgiveness that is still compatible with resisting the invasion of ALLD. For the simplified Prisoner's Dilemma, expressed in the cost and benefit parameters, c and b , we have $q = 1 - c/b$. But the subsequent results suggest revising the definition of GTFT.

3. RESULTS

Figure 1 shows the stationary distribution of our stochastic evolutionary dynamics over the strategy space, which is given by the unit square. Mutants are drawn from a

uniform random distribution. We consider reactive strategies with a minimum noise level: all p and q values are between 0.005 and 0.995. For the three different population sizes, $N = 20, 50$ and 100 , we find that 47, 67 and 81 per cent of the distribution lie in the cooperative half, which is given by $p + q > 1$. Thus, larger population sizes facilitate the evolution of cooperation. We think this is the case, because spite opposes cooperation in very small populations. The stationary distribution has two peaks. One peak is located near the ALLD corner, $(0, 0)$. The second peak is located close to the line $p = 1$. For the three population sizes, the optimum q levels, which correspond to the peak, are given by $q^* = 0.15, 0.23$ and 0.37 , respectively. Larger population sizes favour greater levels of forgiveness. The location of the second peak, $(1, q^*)$, marks the optimum level of forgiveness and therefore serves as a useful definition of GTFT.

The stationary distribution of our evolutionary dynamics represents a mutation–selection balance over a continuous strategy space. The particular shape of the distribution is caused, in part, by the ability of strategies to resist invasion and also by the location of successful invaders. Figure 2 provides the key information. For four strategies, ALLD, TFT, GTFT and ALLC, we show the

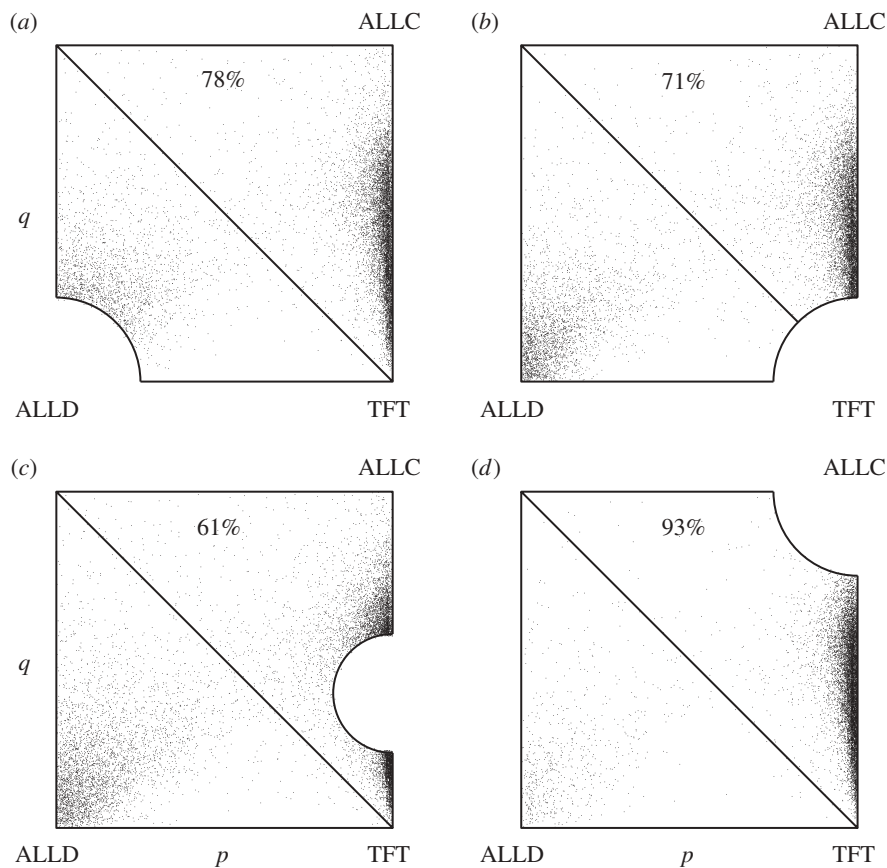


Figure 3. Knock-out experiments. To understand the roles of individual strategies in the evolution of cooperation, we remove certain parts of the strategy space. (a) The point cloud shows the stationary distribution when mutations are drawn from the uniform distribution on all reactive strategies without those close to ALLD. The proportion of time spent in the cooperative half is 78%. (b,c) This proportion drops to 71% if TFT is knocked out instead and to 61% if GTFT with $q = 0.4$ is knocked out. The stationary distribution seems to try to concentrate mass close to the boundary of the removed disc. Here invasion is easier than in the disc, explaining the increase in defection. (d) Knocking out ALLC increases the proportion of cooperation to 93%. These experiments quantify the extent to which TFT and GTFT promote and ALLC hinders the evolution of cooperation. In (a), (b) and (d), the radius of the removed disc is $1/4$, in (c), the radius is $\sqrt{2}/8$, so that the removed areas coincide. The effect of removing strategies is similar but more pronounced if the discs are larger. For example, if a disc with radius $1/3$ around ALLC is removed, then the proportion of cooperation increases to 97%. Parameter values are $N = 100$, $b = 10$ and $c = 1$.

cloud of successful invaders. The arrowhead indicates the centre of gravity of the invading cloud. For ALLD, the first successful mutant appears on average after 319 trials. The arrow points into the vicinity of TFT. Hence, TFT-like strategies are the typical invaders of ALLD. For TFT, however, the first successful mutant appears already after 24 trials on average. Hence, the population remains at TFT only for a short while; TFT is not a peak in the distribution. The arrow points from TFT into the vicinity of GTFT. From there, it takes on average 715 invasion attempts, and the invaders are typically more forgiving. The population moves slowly towards the ALLC corner. For an ALLC population, a successful invader appears on average after 43 attempts. All numerical values are the results of computer simulations for population size $N = 100$; a neutral mutant would need on average 100 attempts for one successful invasion.

The arrows in figure 2 indicate the oscillatory nature of the evolutionary dynamics that is underlying the stationary distribution. From an ALLD population, we move via TFT to GTFT. There, we drift slowly to even more cooperative strategies, which eventually can be invaded by defectors. Then, the cycle starts anew.

In figure 3, we perform ‘knock-out’ experiments, where specific strategies are excluded. We use population size $N = 100$. If the entire strategy space is available for new mutants, then 81 per cent of the stationary distribution is in the cooperative half, $p + q > 1$. If we remove a disc around ALLD, then this value is 78 per cent. If we remove a disc around TFT, then the frequency of cooperative strategies in the stationary distribution drops to 71 per cent. This experiment quantifies the role of TFT in catalysing the emergence of cooperation. Surprisingly, the effect is not very strong. TFT facilitates, but is not needed for the emergence of cooperation. If we knock out a disc around GTFT, which is a peak in the stationary distribution, then the frequency of cooperators falls to 61 per cent. If on the other hand, we knock out a disc around ALLC, then the frequency of cooperative strategies in the stationary distribution increases to 93 per cent. Therefore, ALLC-like strategies are very efficient in catalysing the emergence of defection. They undermine GTFT populations and open the gates for the invasion by defectors. Note that neither TFT nor ALLC are the peaks in the stationary distribution (if the full strategy set is used). Hence, they are not really present

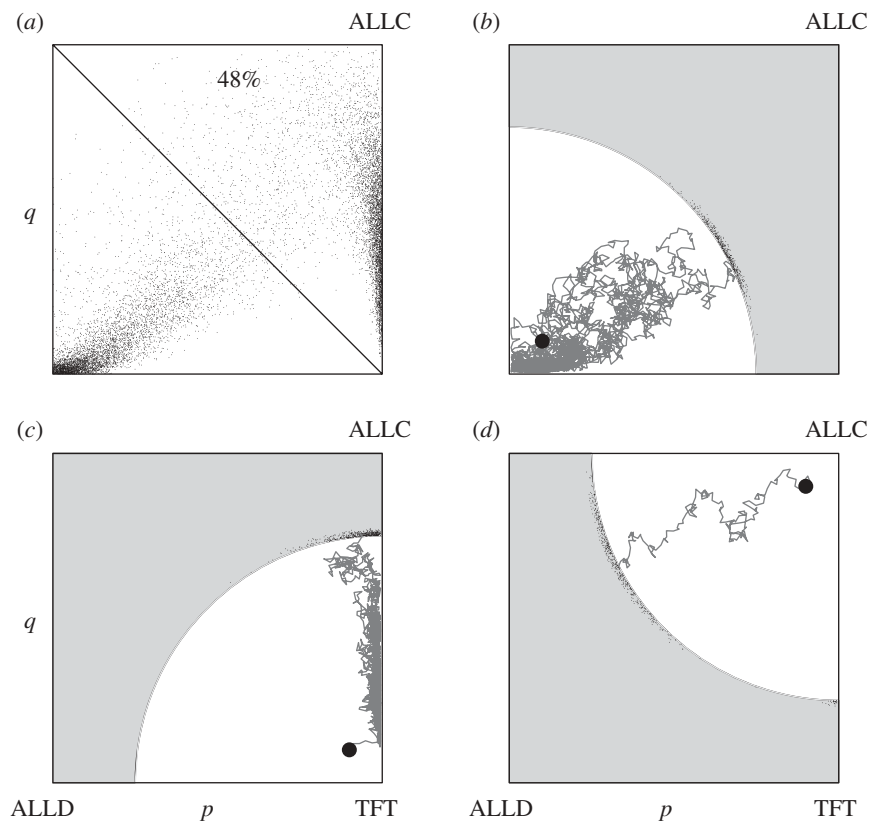


Figure 4. Stochastic evolutionary dynamics with local mutation. In the local version of our stochastic evolutionary dynamics, mutations are drawn from a small neighbourhood around the resident strategy. In these simulations, we use a disc of diameter 0.05. (a) Stationary distribution and fraction of cooperation. As in the global mutation model, the process is most of the time either close to the TFT–ALLC edge or close to the ALLD corner, but the proportion of time spent in the cooperative half is considerably smaller. (b–d) Each panel shows one typical trajectory of the stochastic process starting from the big black point. The small points show the distribution of ‘hitting points’ where the process enters the grey area. They indicate the typical direction of evolutionary dynamics when starting from strategies near ALLD, TFT and ALLC, respectively. Parameter values are $N = 100$, $b = 10$ and $c = 1$.

at equilibrium, but they catalyse the underlying evolutionary dynamics.

(a) Local mutation

So far, we have chosen new mutants from a uniform distribution over the strategy space. We call this approach ‘global mutation’. We can also choose mutants from a local neighbourhood around the resident strategy. We call this approach ‘local mutation’. In this case, long jumps in strategy space are not possible. The evolutionary trajectories have to proceed via small steps. In figure 4a, we show the stationary distribution that is generated by the local mutation model. For a population size of $N = 100$, the frequency of cooperative strategies is about 48 per cent. The peak at ALLD is larger than for the global mutation model; with local mutation, it is harder to replace a population of defectors. In figure 4b–d, we show some evolutionary trajectories of the local mutation model. The distribution of ‘hitting points’ indicates the direction of evolution when starting from certain strategies.

If we choose the invading mutants from an infinitesimally small neighbourhood around the resident strategy, then we obtain a stochastic version of adaptive dynamics (Hofbauer & Sigmund 1990; Nowak & Sigmund 1990; Dieckmann & Law 1996; Metz *et al.* 1996; Dieckmann *et al.* 2000) for finite population size. A system of two

differential equations describes the ‘most probable’ path of evolutionary dynamics. For our strategy space, the deterministic and the stochastic system lead to the same evolutionary trajectories, which are given by concentric quarter cycles around TFT. But interestingly, the direction of evolution can vary for the two approaches. The size of the ‘cooperative region’, where the p and q values increase, increases with population size. The derivation of the stochastic adaptive dynamics is given in the electronic supplementary material.

4. DISCUSSION

We have proposed a new method for studying the stochastic evolutionary dynamics in finite sized populations for games with a continuous strategy space. Our calculation leads to a stationary distribution over the strategy space, which characterizes the relative abundance of strategies in the mutation–selection equilibrium. In contrast to the previous work, our approach does not require weak selection, but we do assume rare mutation: new mutants arrive only one at a time. Our model does not allow different strategies to coexist for long periods of time and is therefore not suitable to study stochastic evolutionary branching.

We have applied our new approach to study the reactive strategies of direct reciprocity. Neither TFT nor

ALLC are successful strategies; they achieve only low abundance in the mutation–selection equilibrium. Knock-out experiments show that TFT is a weak catalyst for promoting cooperation, while ALLC is a strong catalyst for promoting defection. The most successful reactive strategies are ALLD and GTFT. The latter cooperates whenever the other person cooperates, but sometimes cooperates even if the other person has defected. The optimum level of forgiveness depends on the payoff values of the game and the population size. Our approach characterizes the oscillatory nature of the evolution of cooperation. Cooperation is never here to stay. Instead, there are endless cycles between all-out defectors, harsh retaliators, careful forgivers and unconditional cooperators. These cycles reflect to some extent the varying fortunes of human history.

This work was supported by the John Templeton Foundation, the National Science Foundation/National Institutes of Health Joint Program in Mathematical Biology (National Institutes of Health grant R01GM078986) and J. Epstein.

REFERENCES

- Ackermann, M., Stecher, B., Freed, N. E., Songhet, P., Hardt, W. D. & Doebeli, M. 2008 Self-destructive cooperation mediated by phenotypic noise. *Nature* **454**, 987–990. (doi:10.1038/nature07067)
- Axelrod, R. & Hamilton, W. D. 1981 The evolution of cooperation. *Science* **211**, 1390–1396. (doi:10.1126/science.7466396)
- Binmore, K. & Samuelson, L. 1997 Muddling through: noisy equilibrium selection. *J. Econ. Theory* **74**, 235–265. (doi:10.1006/jeth.1996.2255)
- Cressman, R. 2003 *Evolutionary dynamics and extensive form games*. Cambridge, MA: MIT Press.
- Dieckmann, U. & Law, R. 1996 The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* **34**, 579–612. (doi:10.1007/BF02409751)
- Dieckmann, U., Law, R. & Metz, J. A. J. (eds) 2000 *The geometry of ecological interactions: simplifying spatial complexity*. Cambridge, UK: Cambridge University Press.
- Doebeli, M., Hauert, C. & Killingback, T. 2004 The evolutionary origin of cooperators and defectors. *Science* **306**, 859–862. (doi:10.1126/science.1101456)
- Fudenberg, D. & Levine, D. K. 1998 *The theory of learning in games*. Cambridge, MA: MIT Press.
- Fudenberg, D. & Maskin, E. 1990 Evolution and cooperation in noisy repeated games. *Amer. Econ. Rev.* **80**, 274–279.
- Hofbauer, J. & Sigmund, K. 1990 Adaptive dynamics and evolutionary stability. *Appl. Math. Lett.* **3**, 75–79. (doi:10.1016/0893-9659(90)90051-C)
- Hofbauer, J. & Sigmund, K. 1998 *Evolutionary games and population dynamics*. Cambridge, UK: Cambridge University Press.
- Hofbauer, J. & Sigmund, K. 2003 Evolutionary game dynamics. *Bull. Am. Math. Soc.* **40**, 479–519. (doi:10.1090/S0273-0979-03-00988-1)
- Imhof, L. A., Fudenberg, D. & Nowak, M. A. 2005 Evolutionary cycles of cooperation and defection. *Proc. Natl Acad. Sci. USA* **102**, 10 797–10 800. (doi:10.1073/pnas.0502589102)
- Imhof, L. A. & Nowak, M. A. 2006 Evolutionary game dynamics in a Wright–Fisher process. *J. Math. Biol.* **52**, 667–681. (doi:10.1007/s00285-005-0369-8)
- Kandori, M., Mailath, G. & Rob, R. 1993 Learning, mutation, and long run equilibria in games. *Econometrica* **61**, 29–56. (doi:10.2307/2951777)
- Lessard, S. & Ladret, V. 2007 The probability of a single mutant in an exchangeable selection model. *J. Math. Biol.* **54**, 721–744. (doi:10.1007/s00285-007-0069-7)
- May, R. M. 1973 *Stability and complexity in model ecosystems*. Princeton, NJ: Princeton University Press.
- Maynard Smith, J. 1982 *Evolution and the theory of games*. Cambridge, UK: Cambridge University Press.
- Maynard Smith, J. & Price, G. R. 1973 The logic of animal conflict. *Nature* **246**, 15–18. (doi:10.1038/246015a0)
- Metz, J. A. J., Geritz, S. A. H., Meszina, G., Jacobs, F. J. A. & van Heerwaarden, J. S. 1996 Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction. In *Stochastic and spatial structures of dynamical systems* (eds S. J. van Strien & S. M. Verduyn Lunel). *K. Ned. Akad. Van Wet. B* **45**, 183–231. Amsterdam, Holland: North-Holland Publishing Company.
- Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. 2004 Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650. (doi:10.1038/nature02414)
- Nowak, M. A. & Sigmund, K. 1990 The evolution of stochastic strategies in the Prisoner's Dilemma. *Acta Appl. Math.* **20**, 247–265. (doi:10.1007/BF00049570)
- Nowak, M. A. & Sigmund, K. 1992 Tit for tat in heterogeneous populations. *Nature* **355**, 250–253. (doi:10.1038/355250a0)
- Nowak, M. A. & Sigmund, K. 2004 Evolutionary dynamics of biological games. *Science* **303**, 793–799. (doi:10.1126/science.1093411)
- Proulx, S. R. & Day, T. 2001 What can invasion analyses tell us about evolution under stochasticity in finite populations? *Selection* **2**, 1–15.
- Ritzberger, K. & Weibull, J. W. 1995 Evolutionary selection in normal-form games. *Econometrica* **63**, 1371–1399. (doi:10.2307/2171774)
- Samuelson, L. 1997 *Evolutionary games and equilibrium selection*. Cambridge, MA: MIT Press.
- Santos, F. C., Santos, M. D. & Pacheco, J. M. 2008 Social diversity promotes the emergence of cooperation in public goods games. *Nature* **454**, 213–216. (doi:10.1038/nature06940)
- Schaffer, M. 1988 Evolutionarily stable strategies for a finite population and variable contest size. *J. Theor. Biol.* **132**, 469–478. (doi:10.1016/S0022-5193(88)80085-7)
- Tao, Y. & Cressman, R. 2007 Stochastic fluctuations through intrinsic noise in evolutionary game dynamics. *Bull. Math. Biol.* **69**, 1377–1399. (doi:10.1007/s11538-006-9170-0)
- Taylor, P. D. & Jonker, L. B. 1978 Evolutionary stable strategies and game dynamics. *Math. Biosci.* **40**, 145–156. (doi:10.1016/0025-5564(78)90077-9)
- Traulsen, A., Claussen, J. C. & Hauert, C. 2005 Coevolutionary dynamics: from finite to infinite populations. *Phys. Rev. Lett.* **95**, 238701. (doi:10.1103/PhysRevLett.95.238701)
- Trivers, R. L. 1971 The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57. (doi:10.1086/406755)
- Vukov, J. & Szabó, G. 2005 Evolutionary Prisoner's Dilemma game on hierarchical lattices. *Phys. Rev. E* **71**, 036133. (doi:10.1103/PhysRevE.71.036133)
- Weibull, J. W. 1995 *Evolutionary game theory*. Cambridge, MA: MIT Press.