

EVOLUTIONARY DYNAMICS

Christoph Hauert (christoph_hauert@harvard.edu)

Program for Evolutionary Dynamics, Harvard University, One Brattle Square, Cambridge MA 02138, USA

Department of Mathematics, University of British Columbia, 1984 Mathematics Road, Vancouver B.C. Canada V6T 1Z2

Abstract. Evolutionary dynamics in finite populations reflects a balance between Darwinian selection and random drift. For a long time population structures were assumed to leave this balance unaffected provided that the mutants and residents have fixed fitness values. This result indeed holds for a certain (large) class of population structures or graphs. However, other structures can tilt the balance to the extent that either selection is eliminated and drift rules or drift is eliminated and only selection matters.

In nature, however, fitness is generally affected by interactions with other members of the population. This is of particular importance for the evolution of cooperation. The essence of this evolutionary conundrum is captured by social dilemmas: cooperators provide a benefit to the group at some cost to themselves, whereas defectors attempt to exploit the group by reaping the benefits without bearing the costs of cooperation. The most prominent game theoretical models to study this problem are the prisoner's dilemma and the snowdrift game. In the prisoner's dilemma, cooperators are doomed if interactions occur randomly. In structured populations, individuals interact only with their neighbors and cooperators may thrive by aggregating in clusters and thereby reducing exploitation by defectors. In finite populations, a surprisingly simple rule determines whether evolution favors cooperation: $b > ck$ that is, if the benefits b exceed k -times the costs c of cooperation, where k is the (average) number of neighbors. The spatial prisoner's dilemma has led to the general belief that spatial structure is beneficial for cooperation. Interestingly, however, this no longer holds when relaxing the social dilemma and considering the snowdrift game. Due to the less stringent conditions, cooperators persist in populations with random interactions but spatial structure tends to be deleterious and may even eliminate cooperation altogether.

In many biological situations it seems more appropriate to assume a continuous range of cooperative investment levels instead of restricting the analysis to two a priori fixed strategic types. In the continuous prisoner's dilemma cooperative investments gradually decrease and defection dominates just as in the traditional prisoner's dilemma. In contrast, the continuous snowdrift game exhibits rich dynamics but most importantly provides an intriguing natural explanation for phenotypic diversification and the evolutionary origin of cooperators and defectors. Thus, selection may not always favor equal contributions but rather promote states where two distinct types co-exist – those that fully cooperate and those that exploit. In the context of human societies and cultural evolution this could be termed the *Tragedy of the Commune* because differences in contributions to a communal enterprise have significant potential for escalating conflicts on the basis of accepted notions of fairness.

Key words: evolutionary game theory, evolutionary graph theory, social dilemmas, prisoner's dilemma, snowdrift game, structured populations, continuous games, evolutionary branching



1. Modeling Evolution

Evolutionary dynamics acts on populations – neither genes, nor cells, nor individuals evolve, only populations evolve. Conversely, Darwinian selection does not act on populations but on genes, cells and individuals. Selection reflects the fact that the genes or behavioral patterns of individuals with a higher fitness have a higher chance to be passed to subsequent generations through biological reproduction or cultural imitation. For an excellent introduction into evolutionary dynamics see Nowak (2006). In a nutshell, the evolutionary process is determined by:

Selection: Individuals with a fitness that exceeds the average fitness in the population have a higher propensity to pass their genetic or cultural traits to progeny in subsequent generations and these traits are likely to increase in abundance. Similarly, traits that lower the fitness of an individual have small chances to be passed to the next generation, decrease in abundance and eventually disappear.

Variation: Mutations and genetic recombination as well as spontaneous alterations and erroneous imitations of behavioral patterns generate fitness differences among members of the population. Selection acts on these differences and amplifies them over time.

Random drift: In finite populations the transmission of traits through reproduction or imitation is generally a stochastic process. The success of a trait is proportional to the fitness of its carriers but a high fitness does not provide any guarantee for success. With a small probability even the fittest member of the population may not get a chance to pass its trait to the next generation. Similarly, traits of even the least fit individual may get an odd chance to proliferate and persist in the population through random drift. Random drift counteracts selection and becomes increasingly important in smaller populations or for decreasing fitness differences within a population.

Selection, variation and random drift represent the makeup of Darwin's survival of the fittest. It is important, however, to recognize that evolution represents a myopic optimization process. Even though selection always favours individuals with higher fitness, this does not necessarily imply that the average fitness of the population increases. Quite on the contrary, it is often the case that evolution favours traits that reduce the overall fitness of the population. This fundamental problem will become most apparent when addressing the conundrum of the evolution of cooperation in Sect. 3. The reason for such

outcomes lies in the fact that the fitness of an individual is not entirely genetically or culturally predetermined but instead depends on interactions with other members of the population and thus depends on the current abundance and possibly the distribution of traits in the population.

1.1. THE REPLICATOR DYNAMICS

Let us start with the simplest possible evolutionary scenario: an infinitely large population consists of two types, the residents and the mutants. Individuals do not interact and thus the fitness of both types is fixed and independent of their relative abundances. Residents have a normalized fitness of $f_r = 1$ and mutants have a fitness $f_m = r$. If x denotes the fraction of mutants in the population then the evolutionary fate of the mutants is given by the replicator equation (Hofbauer and Sigmund, 1998):

$$\dot{x} = x(f_m - \bar{f}), \quad (1)$$

where $\bar{f} = xf_m + (1-x)f_r$ denotes the average population payoff. This simply means that if $f_m > \bar{f}$ then the mutants increase in abundance. Here this condition reduces to $r > 1$. Thus, if the mutant has a higher fitness than the resident, the fraction of mutants keeps increasing until eventually the mutant displaces the resident and thus becomes the new resident. Conversely, if $r < 1$ the mutant is bound to disappear irrespective of its initial frequency. The evolutionary process is deterministic – if $r > 1$ mutants take over with certainty and disappear with certainty for $r < 1$.

1.2. THE MORAN PROCESS

Evolution in finite populations is stochastic and can be modeled using the Moran process (Moran, 1962): in every time step an individual is randomly selected for reproduction with a probability proportional to its fitness and produces a single clonal offspring that replaces a randomly selected member of the population (see Fig. 1). The total population size N remains constant, i.e. the Moran process assumes that N represents the carrying capacity of the population and neglects fluctuations in population size. All individuals have the same average lifespan but fitter individuals tend to have a higher reproductive output. This represents a specific balance between selection and random drift: fitter individuals have higher chances – but no guarantee – of reproduction, whereas less fit individuals are likely – but again, no guarantee – to be eliminated.

Returning to our simplest possible evolutionary scenario, repeatedly applying the above updating procedure determines the evolutionary fate of residents and mutants. In the absence of mutations the Moran process ultimately

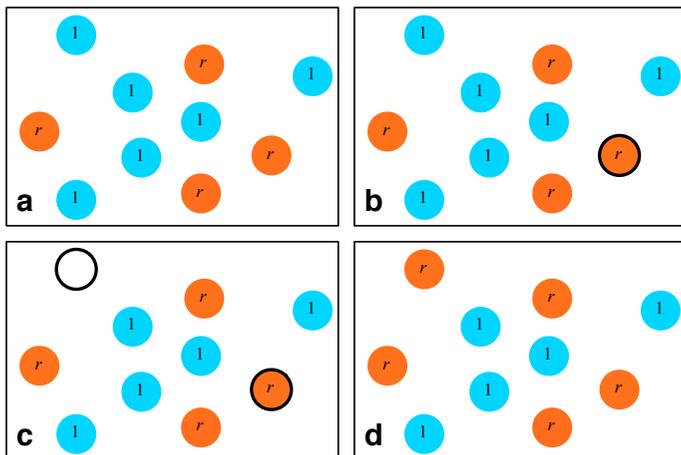


Figure 1. The Moran process describes the stochastic evolution of a finite population of constant size. **a** the population consists of a mixture of residents (blue, fitness 1) and mutants (orange, fitness r). **b** a focal individual is randomly selected for reproduction with a probability proportional to its fitness. **c** a randomly selected individual (independent of fitness) is removed. **d** the vacancy is replaced by a clonal offspring of the focal individual.

results in a homogeneous population with all residents or all mutants because, irrespective of the initial configuration, eventually all members of the population will have a single common ancestor. Both homogeneous states are absorbing and represent an evolutionary end state. The remaining type is said to have reached fixation. We can now determine the probability that mutants (or residents) fixate for a particular initial configuration.

In unstructured populations the state of the population is fully determined by the number of mutants present. The number of mutants i changes at most by ± 1 in every time step of the Moran process. With probability T^+ the number of mutants increases from i to $i+1$, with probability T^- it decreases to $i-1$ and with probability $1 - T^+ - T^-$ the number of mutants remains unchanged.

$$T^+ = \frac{i \cdot r}{i \cdot r + (N - i)} \cdot \frac{N - i}{N} \quad (2a)$$

$$T^- = \frac{N - i}{i \cdot r + (N - i)} \cdot \frac{i}{N} \quad (2b)$$

The first factor of T^+ (T^-) indicates the probability that a mutant (resident) is chosen for reproduction and the second factor denotes the probability that the offspring replaces a resident (mutant). Note that the ratio of the transition probabilities $T^+/T^- = r$ is independent of the number of mutants in the population. This leads to a simple recursive formula for the fixation probability

$\rho(i)$ of the mutant in a population with i mutants:

$$\rho(i) = \frac{r}{1+r} \rho(i-1) + \frac{1}{1+r} \rho(i+1). \quad (3)$$

Thus, the dynamics corresponds to a biased random walk with absorbing boundaries. Eq. (3) admits two solutions $\rho = 1$ and $\rho = 1/r^i$. The absorbing boundaries additionally require $\rho(0) = 0$ and $\rho(N) = 1$. For $r \neq 1$, the fixation probability of a single mutant ρ_1 then becomes

$$\rho_1 = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^N}}. \quad (4)$$

Assuming that mutations are rare events ρ_1 is of particular interest. It is easy to see that a neutral mutant ($r = 1$) has a fixation probability of $\rho_1 = 1/N$: eventually the entire population will have a single common ancestor but in terms of fitness mutants and residents are indistinguishable and so every member of the population has equal chances to be the chosen one. Evolution is said to favor a mutant if the fixation probability of the mutant exceeds the fixation probability of a neutral mutant, $\rho_1 > 1/N$.

In contrast to the replicator dynamics, Eq. (4) shows that evolution favors mutants for $r > 1$ but it is no longer a guarantee to reach fixation and similarly, for $r < 1$ fixation is less likely but again no guarantee that the mutation disappears. The balance between selection and random drift depends on the population size N . For small N random drift dominates whereas for large N selection becomes more important and in the limit $N \rightarrow \infty$ the deterministic replicator dynamics is recovered (Traulsen et al., 2005).

2. Evolutionary Graph Theory

Using evolutionary graph theory (Lieberman et al., 2005) we can investigate the effects of population structures on the evolutionary process. In structured populations, individuals occupy the nodes of a network or graph of size N . The links between the nodes define the neighborhood of each individual. The graph can have any structure – for example, square lattices describe spatially extended systems or small-world networks (Watts and Strogatz, 1998) model social structures – and links between nodes may have different strengths. A fully connected graph, where each node is equally linked to every other node, is equivalent to an unstructured population. Mathematically, the structure of the graph is determined by the adjacency matrix $W = [w_{ij}]$ where w_{ij} denotes the strength of the link between nodes i and j . If $w_{ij} = 0$ and $w_{ji} = 0$ then the two nodes i and j are not connected.

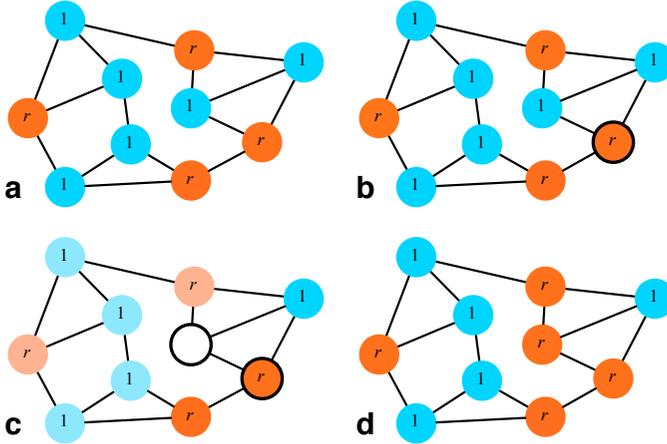


Figure 2. Moran process on graphs – stochastic evolution in a finite structured population of size N . All individuals occupy the nodes of a graph or network where links between nodes determine each individual’s neighborhood. **a** the population consists of a mixture of residents (blue, fitness 1) and mutants (orange, fitness r). **b** a focal individual i is randomly selected for reproduction with a probability proportional to its fitness. **c** a random selected neighbor j of the focal individual is removed. **d** the vacancy is filled by clonal offspring of the focal individual.

The original Moran process is easily adapted to model evolution on arbitrary graphs (see Fig. 2): in every time step a focal individual i is randomly selected for reproduction with a probability proportional to its fitness and produces a single clonal offspring that replaces a random *neighbor* j , which is selected with a probability proportional to w_{ij} .

In structured populations the essential difference is that the offspring replaces a neighbor of the focal individual instead of a random member of the population. An additional minor difference is that the offspring cannot replace the focal individual but, in principle, this can be implemented by adding loops to each node.

Provided that the graph is connected – each node is connected to every other node through a series of links – the system has the same two absorbing states with all residents or all mutants. In the following we always assume connected graphs because otherwise the dynamics must be considered for each subgraph individually. Thus, we can again ask the question about the fixation probability of a single mutant in a structured resident population. In other words, how does the limited dispersal of offspring in structured populations affect the fixation probability of mutants? Some simple sample graphs are shown in Fig. 3. Is fixation easier or harder on these graphs? – or is it independent of the population structure? In fact, the celebrated results by

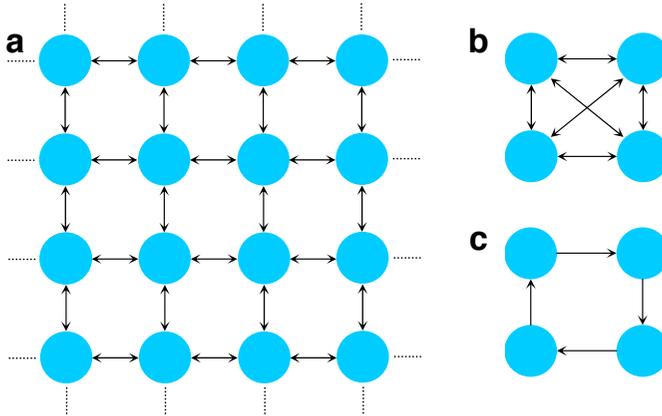


Figure 3. Sample graphs illustrating different simple population structures. What is the fixation probability of a single mutant on each of these graphs? Is it more or less likely than in an unstructured population? **a** regular lattice where each individual has four neighbors. **b** fully connected graph. **c** cyclic structure with directed links – replacement does not need to go both ways.

Maruyama (1970) and Slatkin (1981) indicate that the latter is the case. And indeed, all graphs in Fig. 3 leave the fixation probability unchanged and ρ_1 is the same as in unstructured populations. However, this is not true in general but it does hold for a broad class of graphs.

2.1. CIRCULATION THEOREM

In order to characterize the class of graphs that lead to the same fixation probability ρ_1 as the original Moran process in unstructured populations, let us introduce the flux through each node. The sum of the weights of the incoming links $f_i^{\text{in}} = \sum_j w_{ji}$ denotes the flux entering node i . f_i^{in} relates to a temperature because it indicates the rate at which the occupant of node i gets replaced. 'Hot' nodes are frequently replaced and 'cold' nodes only rarely change their type. In analogy, the sum of the weights of the outgoing links $f_i^{\text{out}} = \sum_j w_{ij}$ denotes the flux leaving node i . f_i^{out} determines the impact of node i on its neighborhood. The matrix W is a circulation if $f_i^{\text{in}} = f_i^{\text{out}}$ holds for all nodes.

Circulation Theorem: The Moran process on a graph results in the same fixation probability ρ_1 of a single mutant as in an unstructured population if and only if the graph is a circulation.

In particular, the circulation theorem holds if W is symmetric ($w_{ij} = w_{ji}$), which holds for all graphs shown in Fig. 3, or if W is isothermal, i.e. all

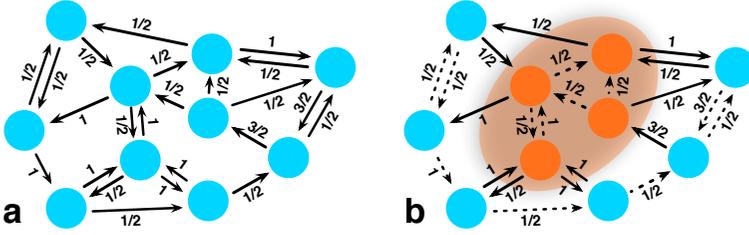


Figure 4. More general graph where mutants have the same fixation probability than in an unstructured population. The graph is a circulation with an asymmetric adjacency matrix W . The arrows indicate the direction and the labels the strength of the links. **a** The flux entering each node is balanced by the flux leaving the node, $f_i^{\text{in}} = f_i^{\text{out}}$. This is the definition of a circulation. **b** mutants (orange) occupy a connected subset of the graph (shaded area) and are surrounded by residents (blue). Note that for the mutant subset holds the same flux balance as for every node: the sum of the weights of incoming links (connecting residents to mutants) is the same as the sum of the weights of outgoing links (connecting mutants to residents). Evolutionary dynamics changes the composition of the population only if replacements occur along one of the solid arrows (connecting the mutant subset and the residents) but not along the dashed arrows. However, the flux balance of the mutant subset is not affected by adding or removing mutants. Because of this invariance mutants have the same fixation probability on a circulation graph as in an unstructured population (see text).

nodes have the same temperature ($f_i^{\text{in}} = T \forall i$). Another special case of the circulation theorem applies if W is doubly stochastic, i.e. if all rows and columns of W sum up to one. A more general circulation graph where W is not symmetric is depicted in Fig. 4a. A detailed proof of the circulation theorem is provided in Lieberman et al. (2005). Here we provide an intuitive illustration of the circulation conditions and its consequences.

At any point in time during the invasion process of mutants on a circulation graph, it is possible to identify connected subsets of nodes on the graph that are occupied by mutants such that all adjacent nodes of each subset are occupied by residents. Obviously, the state of the population changes only if a replacement occurs along one of the links connecting residents and mutants. Figure 4b shows a general circulation graph with one connected subset (shaded area) of mutants. Multiple such subsets may exist and, in fact, the evolutionary process may split large connected subsets of mutants into two smaller ones or may merge two previously unconnected subsets into one larger subset. For each subset, the circulation theorem requires that the sum of the weights of links pointing out of the subset (connecting a mutant node to an adjacent resident node) equals the sum of the weights of links pointing into the subset (connecting an adjacent resident with a mutant node within the

subset). Since the influx is balanced by the outflux, $f_i^{\text{in}} = f_i^{\text{out}}$, for all nodes, increasing the mutant subset by replacing an adjacent resident node with a mutant or decreasing the subset by replacing a mutant with a resident, does not affect the flux balance of the subset. For the same reason, the balance remains unchanged if subsets merge or if one subset splits into two. Recall that the number of mutants in the population changes only if a replacement occurs along any link that connects residents with mutants or vice versa. Because the Moran process essentially selects links with a probability proportional to the link weight and the fitness of the node at its tail, it follows that for each subset, the probability that another mutant is added is simply given by $r/(1+r)$ and the complementary probability that one is removed is $1/(1+r)$. Since this holds for each subset, it also holds for the entire population and is independent of the number, size, shape and distribution of mutant subsets. This invariance applies if and only if the circulation theorem is satisfied. Consequentially, the fixation probability on circulation graphs reduces to the recursive Eq. (3) derived for the original Moran process.

Note that even though the fixation probabilities remain unchanged on circulation graphs, the corresponding fixation times are very sensitive to the details of the population structure and pose a much harder problem. For the original Moran process it can be shown that the fixation time of an advantageous mutant with fitness $r > 1$ in a resident population is the same as the fixation time of a disadvantageous mutant with fitness $1/r$ – however, the first scenario is much more likely (Taylor et al., 2006). The circulation theorem only ensures that the ratio of the transition probabilities $T^+/T^- = r$ remains unchanged but even on circulation graphs T^+ and T^- depend not only on the number but also on the distribution of mutants. Generally, population structures tend to substantially increase the fixation times because the structure limits the possibilities for mutants to conquer new nodes.

The circulation theorem indeed covers a large class of population structures and certainly includes the most intuitive cases such as regular lattices, cycles and fully connected graphs (see Fig. 3). Thus, it is not so surprising that Maruyama (1970) and Slatkin (1981) arrived at the conclusion that introducing population structure leaves fixation probabilities of mutants unaffected. However, the circulation theorem not only allows to determine for which population structures this applies but it also indicates that other structures must exist that *do* result in different fixation probabilities. In particular, what structures suppress selection and enhance random drift ($\rho < \rho_1$ for $r > 1$)? Do they exist and what are their characteristics? And conversely, is it possible to achieve the opposite and generate structures that enhance selection and suppresses random drift ($\rho > \rho_1$ for $r > 1$)?

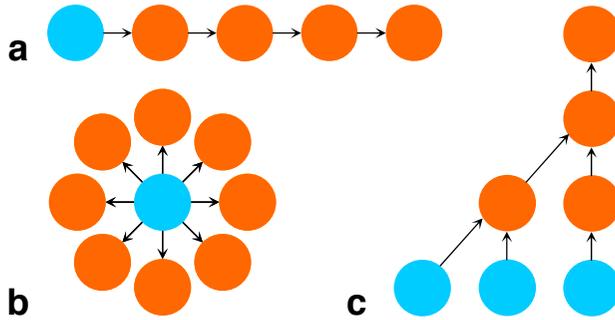


Figure 5. Evolutionary Suppressors are characterized by a hierarchical organization of the population: 'cold' nodes (blue) are infrequently replaced (or not at all) and determine the evolutionary fate of 'hotter' downstream nodes (orange). **a Linear chain:** the offspring of every node replaces the occupant of the node to the right. The leftmost root node is never replaced and the offspring of the rightmost node is lost. **b Burst:** a central hub node feeds into a reservoir. The offspring of the reservoir is lost and the hub is never replaced. The linear chain and the burst structure both result in fixation probabilities of $\rho = 1/N$ – independent of the mutants fitness. A mutant fixates only if it arises in the root or hub node, respectively, but then it fixates with certainty. **c Multiple roots:** In this case no mutant can ever reach fixation, $\rho = 0$. If a mutation occurs in one of the root nodes, it gives rise to a persistent lineage of mutants but it cannot conquer the other root nodes.

2.2. EVOLUTIONARY SUPPRESSORS

Tilting the balance between selection and random drift in favor of random drift means that altering the population structure results in smaller fixation probabilities ρ of an advantageous mutant ($r > 1$) than in the original Moran process or on a circulation graph, $\rho < \rho_1$. Conversely it means that a disadvantageous mutant ($r < 1$) has a higher fixation probability, $\rho > \rho_1$. Indeed, this holds whenever a population is arranged in a hierarchical manner. Some examples of such evolutionary suppressors are shown in Fig. 5. The most extreme case is given by a linear chain where the offspring of each individual replaces the occupant of the node to its right (see Fig. 5a). The leftmost node is a root node and is never replaced whereas the offspring of the rightmost node is lost. This generates a flux through the population from left to right such that no mutant can reach fixation unless the mutation occurs in the root node. This happens with the probability $1/N$ but in that case fixation occurs with certainty. Thus the fixation is simply $\rho = 1/N$, irrespective of the mutant's fitness and hence selection is eliminated and random drift rules. If there are multiple roots (see Fig. 5c) no mutation can ever fixate, $\rho = 0$.

Evolutionary suppressors have a very simple, almost trivial structure but at the same time they turn out to be highly relevant in biological systems.

While mutations enable populations to adapt to changing environments, they are generally pathogenic when they occur within an organism. Especially dangerous are those mutations that increase the net reproductive rate of a cell because this may later develop into cancer (Vogelstein and Kinzler, 1998). In order to prevent accumulation and spreading of detrimental mutations, organisms take advantage of evolutionary suppressors (Nowak et al., 2003). Epithelial tissue, such as our skin or the colon, is organized into small compartments (crypts in the colon) and each compartment is arranged in multiple layers of cells of increasing degrees of differentiation – ranging from few undifferentiated stem cells to terminally differentiated epithelial cells. With the exception of the stem cells, all cells are regularly renewed by new cells from precursor layers. This exactly matches the setup of the linear chain (see Fig. 5a) and thus cancerous mutations will be eventually washed out, unless they happen to occur in one of the stem cells. In case this occurs, then the compartmentalization confines the mutants and prevents further spreading. Another impressive example of a complex hierarchical arrangement is given by our blood system where the stem cells reside in the bone marrow and divide only about once a week through a series of precursor cells to the terminally differentiated red blood cells with a production of the order of 10^{12} cells every day. The architecture of all these systems is shaped to prevent malignant mutations from spreading.

2.3. EVOLUTIONARY AMPLIFIERS

Evolutionary amplifiers are the counterpart to evolutionary suppressors. These population structures tilt the balance between selection and random drift in favor selection such that the fixation probabilities ρ of advantageous mutants ($r > 1$) is larger than in the original Moran process or on circulation graphs, $\rho > \rho_1$. Because selection is enhanced, this also implies that disadvantageous mutants ($r < 1$) have a smaller fixation probability, $\rho < \rho_1$. Evolutionary amplifiers are also characterized by hierarchical population structures with the crucial addition of positive feedback loops. A selection of evolutionary amplifiers is shown in Fig. 6. The simplest example is given by the *star* structure where a central hub is connected to a reservoir of leaf nodes (see Fig. 6a) but in contrast to the *burst* structure (see Fig. 5b) the links between the hub and the leaves are bi-directional. The hub represents a bottleneck for the evolutionary progression because if one leaf node is occupied by a mutant it needs to conquer the hub before another leaf node can be taken over. Most of the time, the 'hot' hub is replaced by reproducing leaf nodes and only occasionally the hub itself reproduces and replaces a leaf node. For an advantageous mutant in a leaf node this means that compared to a resident leaf node it has a relative advantage of r to occupy the hub and similarly the

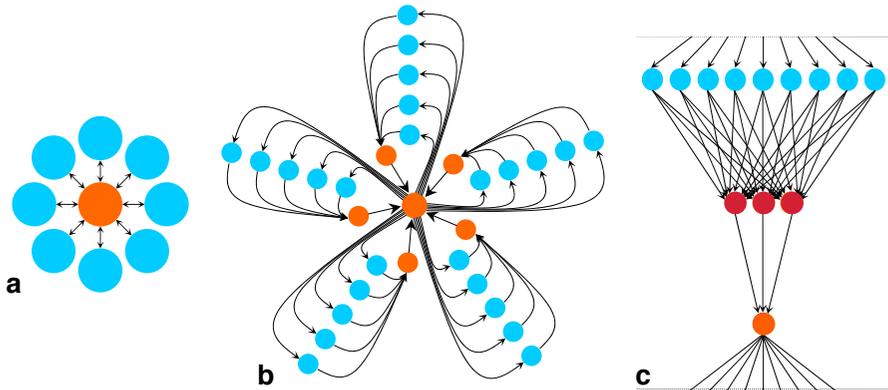


Figure 6. Evolutionary Amplifiers are characterized by a hierarchical organization of the population where 'cold' reservoir nodes (blue) feed through a series of bottlenecks into 'hotter' (orange, red) nodes and eventually into a central hub that feeds back into the reservoir. The diameter k of the graph, i.e. the minimum number of links that connect any node with any other node, determines the amplification of the graph. A mutant with fitness r fixates with the same probability as another mutant with fitness r^k on a circulation graph. **a Star**: a central hub node is connected to a reservoir of leaf nodes, $k = 2$. **b Superstar**: several petals consist of a reservoir that is connected through a linear chain to a central hub that feeds back into all reservoirs, $k = 3$. **c Funnel**: a reservoir feeds into a smaller downstream sub-population that in turn feeds into a hub, which then feeds back into the reservoir, $k = 3$.

mutant hub has again a relative reproductive advantage of r . Thus, the overall relative advantage of a mutant leaf node to proliferate and occupy another leaf node is r^2 . Note that there is no other way for a mutant to spread through the population. As we will see, a mutant on the star structure with fitness r has the same fixation probability as a mutant with fitness r^2 on a circulation graph.

The principles of the star structure can be generalized to a series of sequential bottlenecks that further increase amplification such as provided by the *superstar* (Fig. 6b) or the *funnel* (Fig. 6c). The decisive quantity in all amplifiers is the minimum number of reproductive steps k that are required for a 'cold' reservoir node to populate another reservoir node. k determines the strength of the evolutionary amplifier. In an unstructured population, every node can be reached in one step and thus $k = 1$, on the star it takes two steps, $k = 2$, and on the superstar and funnel shown in Fig. 6 it takes three steps, $k = 3$. The latter two structures can be easily generalized to arbitrary k and the fixation probability becomes (Lieberman et al., 2005):

$$\rho_k = \frac{1 - \frac{1}{r^k}}{1 - \frac{1}{r^{kN}}}. \quad (5)$$

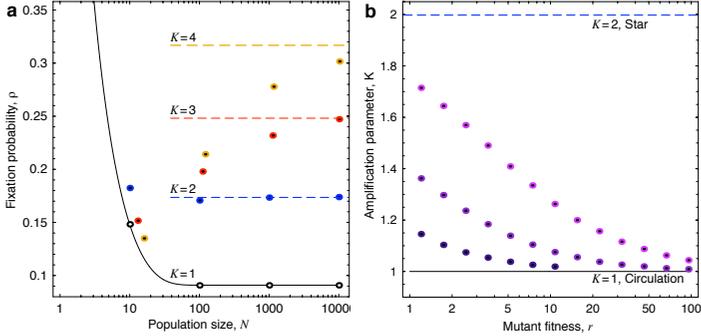


Figure 7. Simulation data for evolutionary amplifiers. **a** fixation probabilities for an advantageous mutant with fitness $r = 1.1$ for different population structures as a function of the population size N : circulation graph (black, $k = 1$), star (blue, $k = 2$) and superstars (red, $k = 3$; yellow $k = 4$). The solid black line indicates the theoretical prediction for circulations and the dashed color lines denote the prediction in the limit $N \rightarrow \infty$. **b** strength of amplification on scale-free networks of size $N = 100$ as a function of the mutant fitness r . The amplification decreases with increasing fitness and with increasing average connectivity m of the nodes: $m = 2$ (purple), $m = 4$ (violet) and $m = 6$ (navy). For marginally beneficial mutations and $m = 2$ the amplification is almost as strong as on a star structure but then gradually declines and approaches $k = 1$ for circulation graphs.

Thus, the fixation probability ρ_k of a mutant with fitness r on an evolutionary amplifier with strength k is the same as that of a mutant with fitness r^k in the original Moran process or on a circulation graph. In finite populations Eq. (5) is only an approximation and becomes exact for $N \rightarrow \infty$. Comparisons with simulation data is shown in Fig. 7a. The proof of Eq. (5) involves a recursive analysis of the sequential bottlenecks as sketched above for the star structure. Technical details are provided in Lieberman et al. (2005).

In the limit $N, k \rightarrow \infty$ the fixation probability even for a mutant with an arbitrarily small fitness advantage converges to one. Similarly, the elimination of a mutant with only an arbitrarily small fitness disadvantage happens with certainty. However, evolutionary amplification comes at a price, which is reflected in long fixation times. In fact, fixation times tend to infinity as fixation probability approaches certainty.

In contrast to evolutionary suppressors, population structures that amplify selection are rather complex and potentially less relevant in nature. However, one interesting exception occurs in the case of scale-free networks (Albert and Barabási, 2002). These networks are characterized by a power law degree distribution, where the degree or connectivity of a node indicates its number of neighbors. Therefore, few nodes are highly connected whereas most nodes entertain only few connections to other nodes. This seems to capture relevant features of social, technological and biological systems ranging from the net-

work of scientific collaborations, to U.S. power lines (Barabási and Albert, 1999) and gene regulatory networks (Koonin et al., 2006).

On scale-free networks the amplification factor k depends on the fitness of the mutant r (see Fig. 7b) and decreases with increasing fitness. Thus, scale-free networks selectively support mutants that are only marginally advantageous – such mutations are prone to accidental extinction through random drift – but the structure does not promote highly beneficial mutations – these fixate with very high probability anyways.

3. The Problem of Cooperation

In nature the fitness of individuals is, in general, not fixed as we have assumed so far, but instead depends on interactions with other members of the population. Evolutionary game theory (Maynard Smith, 1982; Maynard Smith and Price, 1973) provides a powerful mathematical framework to analyze situations where the performance of an individual does not only depend on its own behavior but also on the behavior of its interaction partner or opponent. The most interesting scenario refers to the evolutionary puzzle of the emergence of cooperation under Darwinian selection. The problem of cooperation is captured by social dilemmas (Dawes, 1980), which describe a conflict of interest between the community and the individual. In social dilemmas cooperators produce a valuable public good at some cost to themselves while defectors attempt to exploit the common resource without contributing themselves. Because the public good is valuable, groups of cooperators are better off than groups of defectors, but in any mixed group defectors outperform cooperators – and hence the dilemma. The most prominent game theoretical models to investigate this kind of interactions are the prisoner's dilemma and the snowdrift game for pairwise interactions as well as the public goods game for interactions in larger groups (Doebeli and Hauert, 2005). Here we focus on pairwise interactions in the prisoner's dilemma and snowdrift game and refer to Hauert et al. (2006c) for a general discussion of social dilemmas in groups of arbitrary size.

Social dilemmas are abundant in nature. For example, musk oxen create defense formations to protect their young from wolves (Hamilton, 1971). However, for each ox it would be better to avoid potential injury and to stand in the second line but if every individual behaves that way their defense breaks down and the group becomes prone to attacks by wolves. Similar conflicts of interests occur in sentinel behavior in meerkats (Clutton-Brock et al., 1999), in predator inspection behavior in fish (Milinski, 1987; Pitcher, 1992), in phages competing for reproduction (Turner and Chao, 1999; Turner and Chao, 2003) or in microorganisms producing extra cellular products such as

enzymes in yeast (Greig and Travisano, 2004), biofilms (Rainey and Rainey, 2003) or antibiotic resistance (Neu, 1992), to name only a few prominent examples. However, social dilemmas also occurred on an evolutionary scale and life could not have evolved without the repeated incorporation of lower level units into higher level entities. Major transitions such as the formation of chromosomes out of replicating DNA molecules, the transition from single cells to multicellular organisms or from individuals to societies require cooperation (Maynard Smith and Szathmary, 1995). Finally, humans have taken the problem of cooperation to yet another level (Hardin, 1968) when it comes to social security such as health care or pension plans and, even more importantly, to global scales in terms of natural resources such as drinking water, clean air, fisheries and climate (Milinski et al., 2006).

3.1. PRISONER’S DILEMMA

The prisoner’s dilemma made its first appearance in an experimental bargaining setup designed Melvin Dresher and Merrill Flood in the wake of the second world war (Flood, 1958). Only later it was named by Albert W. Tucker, who contributed an illustrative anecdotal story where two burglars are arrested on the suspicion of a robbery (Poundstone, 1992). Each burglar is interrogated separately and has the options to either refuse to give evidence or to blame his fellow prisoner. They both know that if both refuse to give evidence they will be charged for a minor crime and sentenced to one year imprisonment but if they blame each other, they face three years imprisonment. However, if one refuses to give evidence but gets blamed by the other, then the first one gets the full charge of five years whereas the approver is set free. It is easy to see that no matter what the fellow prisoner decides it is always better to blame him or her in order to reduce the sentence. But if both prisoners reason the same way, they both end up three years in prison instead of only one if they had refused to give evidence. Thus, selfish interests prevent them from achieving the mutually preferable outcome. A state where none of the participants can improve its payoff by unilaterally changing its strategy is called a Nash equilibrium (Nash, 1951).

In evolutionary biology, the prisoner’s dilemma is usually framed in terms of fitness costs and benefits. Cooperators provide a benefit b to their co-player at a cost c to themselves ($b > c$) and defectors neither provide benefits nor pay costs. The payoffs for the joint behavior of two interacting individuals is usually written in the form of a payoff matrix:

$$\begin{array}{cc}
 & \begin{array}{cc} C & D \end{array} \\
 \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}.
 \end{array} \tag{6}$$

Mutual cooperation pays $b - c$ whereas mutual defection pays nothing. However if only one player defects and the other cooperates then the defector gets the benefit b without having to pay the costs and the cooperators faces the costs c without receiving any benefit. Thus, just as in the case of the prisoners, it is always better to defect irrespective of the other players behavior but if both players follow this reasoning they end up with nothing instead of $b - c$.

Evolutionary dynamics is about populations and in this case about the change in frequencies of cooperators and defectors. In an infinite population with a fraction x cooperators ($1 - x$ defectors) and randomly interacting individuals, the evolutionary fate of cooperators is given by the replicator equation:

$$\dot{x} = x(f_C - \bar{f}) = x(1 - x)(f_C - f_D), \quad (7)$$

where f_C, f_D represent the average payoffs of cooperators and defectors, respectively, and $\bar{f} = xf_C + (1 - x)f_D$ denotes the average population payoff. The average payoff of cooperators is simply $f_C = xb - c$ because in every interaction they pay the costs of cooperation c but only if they meet another cooperator they receive the benefit b , which happens with probability x . Similarly the average payoff of defectors is $f_D = xb$. Thus, cooperators are always worse off ($f_C < f_D$) and irrespective of their initial frequency, they will dwindle and eventually disappear. $x^* = 0$ is the only stable equilibrium. This nicely illustrates the fact that evolutionary dynamics represents a myopic optimization process: even though fitter individuals are selected in every time step, the overall fitness of the population decreases.

In finite populations, the fitness of a player is given as $1 - w + wP$, i.e. the convex combination of a baseline fitness, which is normalized to 1 for all players, and the payoff P from the prisoner's dilemma interactions. The relative importance of the two components is specified by w . For $w \rightarrow 0$, fitness differences decrease and selection becomes weak. In order to model evolution, the Moran process is equally applicable to settings where the fitness depends on the current composition of the population, i.e. if fitness is frequency dependent (Nowak et al., 2004). In a population of size N with i cooperators, the average fitness of cooperators and defectors is given by:

$$f_C(i) = 1 - w + \frac{w}{N - 1} ((i - 1)b - (N - 1)c) \quad (8a)$$

$$f_D(i) = 1 - w + \frac{w}{N - 1} ib. \quad (8b)$$

Note that for the Moran process $f_C(i), f_D(i) > 0$ must always hold in order to translate fitness into probabilities of reproduction but this is easily achieved by limiting the maximum selection strength. The replicator equation does not impose similar constraints because the fitness denotes the rate of reproduction relative to the population average.

Based on Eq. (8) the transition probabilities T^+ , T^- for a change to $i + 1$ or $i - 1$ cooperators can be derived in analogy to Sect. 1.2. However, solving the recursive equation in order to determine the fixation probability ρ_C of a single cooperator in a population of defectors is a bit more challenging (Nowak et al., 2004; Karlin and Taylor, 1975):

$$\rho_C = 1 / \left(1 + \sum_{k=1}^{N-1} \prod_{i=1}^k \frac{f_D(i)}{f_C(i)} \right). \quad (9)$$

In the weak selection limit, $w \rightarrow 0$, Eq. (9) simplifies to

$$\rho_C = \frac{1}{N} - \frac{w}{2N}(b + (N - 1)c) + o(w^2) < \frac{1}{N}. \quad (10)$$

and hence evolution never favors cooperation – cooperators are doomed just as in infinite populations.

3.2. SNOWDRIFT GAME

The anecdotal story behind the snowdrift game involves two drivers on their way home caught in a blizzard and trapped on either side of a snowdrift (Sugden, 1986). Each driver has the option to remove the snowdrift and start shoveling or to remain in the cozy warmth of the car. If both cooperate and shovel, they both receive the benefit b of getting home while sharing the labour costs c ($b > c$) but if only one shovels both still get home but the cooperator has to do all the work. If no one shovels neither gets anywhere and they have to wait for spring to melt the snowdrift or at least for the rescue team. In contrast to the prisoner's dilemma, the best strategy now depends on the co-player's decision: if the other driver shovels it is best to be lazy but when facing a lazy bum it is better to swallow the bitter pill and to start shoveling instead of remaining stuck in the snow.

The payoff matrix of the snowdrift game is given by

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \left(\begin{array}{cc} b - \frac{c}{2} & b - c \\ b & 0 \end{array} \right). \end{array} \quad (11)$$

The snowdrift game has the same characteristics as the game of Chicken or the Hawk-Dove game (Maynard Smith, 1982) but these games are usually framed in terms of competitive interactions. Note that for $2b > c > b$ the snowdrift game turns into a prisoner's dilemma. For even higher costs, $c > 2b$ mutual defection becomes the mutually preferred outcome.

In infinite populations with a fraction x cooperators the evolutionary dynamics is again determined by the replicator equation (7) with $f_C = b -$

$c(1 - x/2)$ and $f_D = xb$. In contrast to the prisoner's dilemma, $x^* = 0$ is now unstable and an interior fixed point exists with $f_C = f_D$ for $x^* = 1 - r$ and $r = c/(2b - c)$. Thus, in the snowdrift game cooperators and defectors can co-exist at some equilibrium frequency, which is determined by the costs and benefits of the game. This originates in the fact that in the snowdrift game it is always better to adopt a strategy that differs from the co-player. As a consequence $f_C > f_D$ holds if cooperators are rare ($x \rightarrow 0$) but $f_C < f_D$ if cooperators abound and defectors are rare ($x \rightarrow 1$). Note that at the equilibrium, the population as a whole is worse off than if everybody would cooperate ($\bar{f} = (1 - r)^2(b - c/2) < b - c/2$) – this is the hallmark of social dilemmas and another instance of myopic optimization.

In finite populations we could proceed as before and determine the fixation probability of a single cooperator in a defector population. However, because cooperators and defectors can co-exist, the fixation probability may no longer be a relevant quantity to characterize the dynamics of this system. In fact, co-existence easily leads to exceedingly long fixation times and so situations are possible where fixation of cooperation is highly likely but requires eons to happen. Thus, with the exception of extremely small populations, equilibrium properties seem more relevant to characterize the snowdrift game in infinite as well as in finite populations.

3.3. SPATIAL GAMES

In the prisoner's dilemma cooperators are doomed in the absence of supporting mechanisms. Over the last decades several mechanisms have been proposed that are capable of establishing and maintaining cooperation among unrelated individuals. The different mechanisms essentially fall into four categories: (i) conditional behavioral rules under direct or indirect reciprocity (Trivers, 1971; Nowak and Sigmund, 1998); (ii) extensions of the strategy space by allowing for punishment or voluntary participation (Clutton-Brock and Parker, 1995; Hauert et al., 2002; Hauert et al., 2007); (iii) feedback between ecological and evolutionary dynamics (Hauert et al., 2006a; Hauert et al., 2006b); or (iv) by introducing population structures (Nowak and May, 1992; Hauert and Doebeli, 2004; Ohtsuki et al., 2006).

In this section we consider the last case and return to structured populations where individuals occupy nodes on a graph (c.f. Sect. 2) such that individuals no longer interact with all members of the population. The links of the graph define the neighborhood of all individuals and their fitness is based on interactions within this local neighborhood. As before, we are interested in how population structure affects the evolutionary dynamics and the fate of cooperators and defectors, in particular. Unfortunately, this is a hard problem and is analytically intractable in general because the fitness of each

individual depends on the local configuration of its neighborhood. In fact, to fully understand the evolutionary dynamics in structured populations can be even challenging for computer simulations. Despite these bleak perspectives, there are interesting and relevant exceptions that reveal new insights into the interplay of cooperation and spatial structure.

Among the most important results ranks the finding that spatial structure supports cooperation in the prisoner's dilemma (Nowak and May, 1992). In the spatial prisoner's dilemma a cooperator provides a benefit b to all of its k neighbors at a cost kc to itself. Defectors do not provide benefits and pay no costs. Thus, a cooperator with k neighbors and i cooperators among them has an average fitness of $f_C^i = 1 - w + w(ib - kc)/k$ and a defector in the same position achieves $f_D^i = 1 - w + wib/k$ where w specifies the selection pressure on the performance in the prisoner's dilemma (c.f. Eq. (8)). Structured populations enable cooperators to thrive by forming clusters and thereby they more often interact with other cooperators and, at the same time, they reduce exploitation by defectors. However, this clustering advantage is limited and requires sufficiently small cost-to-benefit ratios c/b .

There are two fundamentally different approaches to investigate effects of space on cooperation: first we consider finite populations of size N and determine the conditions under which spatial structure promotes the evolution of cooperation such that a single cooperator in a defector population has a higher fixation probability than a neutral mutant, $\rho_C > 1/N$. In the next section a surprisingly simple and general rule is derived in the limit of weak selection, $w \rightarrow 0$, based on pair approximation (Matsuda et al., 1992; van Baalen and Rand, 1998). In Sect. 3.5 we turn to equilibrium properties in situations where cooperators and defectors co-exist for long times. In particular, it turns out that space affects cooperation rather differently in the prisoner's dilemma and the snowdrift game.

3.4. THE $B > C \cdot K$ -RULE

In order to derive the fixation probability of a single cooperator ρ_C , the evolutionary dynamics of cooperators and defectors on a graph can again be modeled by the Moran process where the fitness of each individual depends on interactions with all other individuals in its neighborhood (Ohtsuki et al., 2006). A sample graph with prisoner's dilemma fitness values is shown in Fig. 8a. The original Moran process is formulated as a death-birth process (see Sect. 1.2): an individual is selected for reproduction with a probability proportional to its fitness and its offspring replaces a random member of the population (on graphs a random neighbor is replaced). However, the sequence of events could be easily reversed such that first a random member of the pop-

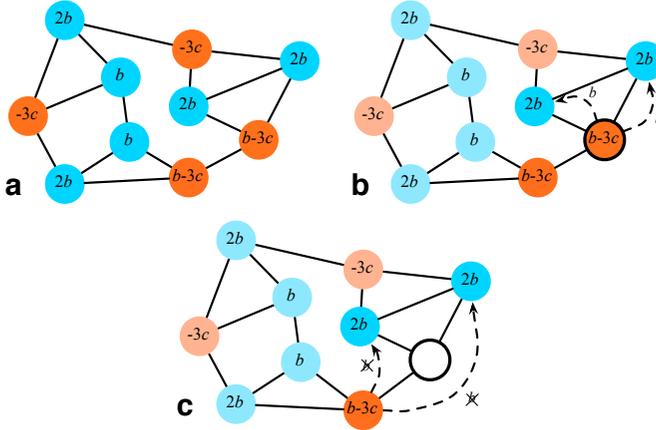


Figure 8. Games on graphs – cooperators (orange) and (defectors) compete on a regular graph where each node has $k = 3$ neighbors. **a** *local interactions* – for the prisoner’s dilemma, each node is marked with the fitness contribution arising from interactions with the three neighbors. **b** *birth-death updating* – the bordered cooperator was selected for reproduction and its offspring will replace one of its three neighbors. Based on their fitness, neighbors compete for reproduction but this puts cooperators at a disadvantage because they supported their defecting competitors by providing them with a benefit b (dashed arrows). **c** *death-birth updating* – the bordered node became vacant and its neighbors compete to repopulate the node based on their fitness. In this case cooperators are better off because less or no support goes to the defecting competitors (dashed arrows).

ulation is removed and then the remaining individuals compete to repopulate the vacant site (on graphs only the neighbors of the vacant site compete). In unstructured populations or if fitness is fixed, changing the sequence of events manifests itself in only marginal changes of the results. However, in the current setup it turns out to be of crucial importance. In fact, for the birth-death process structured populations are unable to promote cooperation and $\rho_C < 1/N$ always holds (Ohtsuki et al., 2006). The intuitive reason is that neighboring individuals compete for reproduction, which means that cooperators actually support their defecting competitors (see Fig. 8b). In contrast, for the death-birth process the disadvantage of cooperators is reduced because the individuals competing to repopulate the vacant site are typically not neighbors (even though they can be) and thus cooperators rarely feed their competitors (see Fig. 8c). Indeed, we shall see that for the death-birth process, evolution can favor cooperation if $b > c$ holds.

3.4.1. *Pair Approximation*

Analytical approximations of the fixation probability ρ_C are based on pair approximation, which requires *regular* or *homogeneous* graphs. On homogeneous graphs each individual has the same number of neighbors k , all links are undirected and have identical weights. Thus, the graph looks the same when viewed from any one node. This holds for any lattice (see Fig. 3 for a square lattice) as well as for less uniform structures such as random regular graphs (Bollobás, 1995) – an example with $k = 3$ is shown in Fig. 8a.

In unstructured populations the evolutionary change is described by the change in the number or frequency of cooperators (see Sect. 3.1). In structured populations this is insufficient because it neglects local correlations but it is possible to track nearest neighbor correlations using pair approximation (Matsuda et al., 1992; van Baalen and Rand, 1998). For homogeneous graphs, the pair configuration probability or the frequency of a strategy pair $p_{ss'}$ indicates the chance that when randomly picking an individual and one of its neighbors that the individual has strategy s and the neighbor s' . For our purposes s, s' are either cooperation c or defection d . p_s is simply the frequency of strategy s with $p_s = p_{sc} + p_{sd}$ and because $p_c + p_d = 1$ it follows that $p_{cd} = p_{dc}$. Configuration probabilities of arbitrarily large clusters are approximated by pair configuration probabilities. For example, the probability of the three-cluster s, s', s'' is given by $p_{ss's''} = p_{ss'} \cdot p_{s's''} / p_{s'}$ where the denominator corrects for the fact that both, $p_{ss'}$ and $p_{s's''}$ include the probability for s' . Note that this approximation applies to tree graphs (or Bethe lattices) but neglects loops which are particularly important on lattices. Note that pair approximation cannot distinguish any graphs with the same k . Since random regular graphs are locally similar to trees (Bollobás, 1995), predictions based on pair approximation are expected to be better for random regular graphs than for lattices.

The spatial dynamics can thus be approximated by four variables p_{cc} , p_{cd} , p_{dc} and p_{dd} but because they must add up to 1 and because of $p_{cd} = p_{dc}$ this only requires two dynamical equations. The most interesting quantities are the overall fraction of cooperators $p_c = p_{cc} + p_{cd}$ as well as the local affinity of cooperators, i.e. the conditional probability that the neighbor of a cooperator is another cooperator $q_{c|c} = p_{cc}/p_c$. In the following the dynamics is expressed in terms of these two quantities but to keep the formulas simple, another local quantity is sometimes used: $q_{c|d} = p_{cd}/p_d = (1 - p_c(2 - q_{c|c})) / (1 - p_c)$, i.e. the conditional probability that the neighbor of a defector is a cooperator.

3.4.2. Dynamical equations

The change in p_c and $q_{c|c}$ is determined by the evolutionary dynamics and in this case the Moran process. Thus we need to determine the probability that a cooperator is replaced by a defector (or vice versa) as well as the effects of such a replacement on p_c and $q_{c|c}$. In the death-birth process, if a defector was eliminated, the neighborhood of the vacant site consist of k_c cooperators with probability

$$\binom{k}{k_c} q_{c|d}^{k_c} (1 - q_{c|d})^{k - k_c}. \quad (12)$$

The neighboring cooperators and defectors have an average fitness of

$$f_c = 1 - w + w((k - 1)q_{c|c} b - k c) \quad (13a)$$

$$f_d = 1 - w + w((k - 1)q_{c|d} b). \quad (13b)$$

Note that each neighbor had at least one defector (the now vacant site) in its own neighborhood. Thus, the probability that the offspring of a cooperator succeeds in repopulating the vacant site becomes

$$\frac{k_c f_c}{k_c f_c + (k - k_c) f_d}. \quad (14)$$

If the defector is replaced by a cooperator, this increments p_c by $1/N$ and $q_{c|c}$ by $2k_c/(kN)$. The total increments are given by the sum over all $k_c = 0, \dots, k$. Decrements in p_c and $q_{c|c}$ arise from replacing a cooperator by a defector and are given by an analogous calculation. After some algebra we obtain the leading terms in w of the dynamical equations:

$$\dot{p}_c = w \cdot \frac{k - 1}{kN} p_c (1 - q_{c|c}) (1 + q_{c|c} - q_{c|d}) \times \\ (b(k - 1)(q_{c|c} - q_{c|d}) - c k) + O(w^2) \quad (15a)$$

$$\dot{q}_{c|c} = \frac{2}{kN} (1 - q_{c|c}) \left[1 - (k - 1)(q_{c|c} - q_{c|d}) \right] + O(w). \quad (15b)$$

Detailed derivations are provided in Ohtsuki et al. (2006). Eq. (15) cannot be solved analytically in general but in the weak selection limit, $w \rightarrow 0$, a natural separation of time scales occurs where $q_{c|c}$ equilibrates much more quickly than p_c and thus the dynamical system rapidly converges to the slow manifold defined by $\dot{q}_{c|c} = 0$, or more explicitly by $q_{c|c} - q_{c|d} = 1/(k - 1)$. This yields

$$q_{c|c} = p_c + \frac{1}{k - 1} (1 - p_c) \quad (16)$$

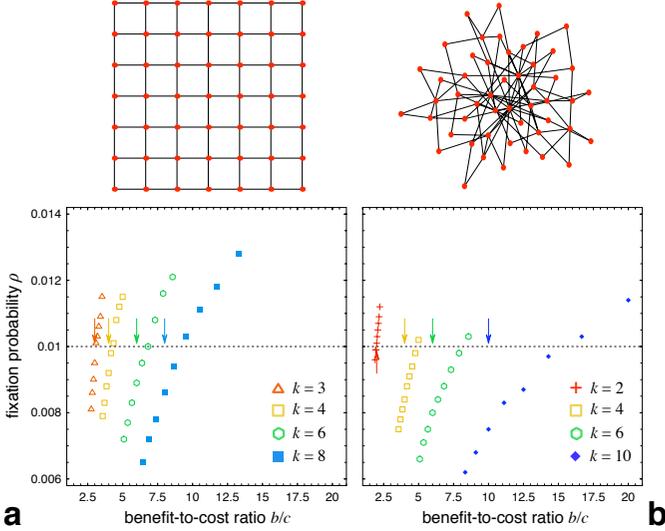


Figure 9. Fixation probability ρ_C of a single cooperator on different types of graphs of size $N = 100$ as a function of the benefit to cost ratio b/c . The dotted horizontal line marks the fixation probability of a neutral mutant ($1/N = 0.01$) and the arrows indicate the predictions by the $b > c k$ -rule. **a regular lattices:** Deviations increase for larger k because pair approximation is based on the assumption that $N \gg k$ holds. On top, a sample lattice with $k = 4$ is shown. **b scale-free networks:** good predictions are obtained even for highly inhomogeneous graphs that violate the assumptions of pair approximation. The networks are generated according to preferential attachment (Albert and Barabási, 2002) and a sample network with an average connectivity of $\langle k \rangle = 4$ is shown on the top.

and, upon neglecting higher order terms in w , the dynamics on the slow manifold becomes

$$\dot{p}_c = w \cdot \frac{k-2}{(k-1)N} p_c (1-p_c) (b-c k) \quad (17)$$

Thus, in the weak selection limit, the fraction of cooperators increases provided that $b > c k$ holds. In order to derive the fixation probability ρ_C we assume that Eq. (16) always holds. This allows to consider a diffusion process of the random variable p_c on the slow manifold. Determining the drift and variance of the diffusion process leads to a backward Kolmogorov differential equation for the fixation probability of a single mutant with the solution

$$\rho_C = \frac{1}{N} + w \frac{N-1}{2N} (b - k c). \quad (18)$$

It follows that $\rho_C > 1/N$ holds if and only if $b > c k$. This result is confirmed by extensive numerical simulations on various kinds of graphs (see Fig. 9). The

simulations clearly show that the condition $b > ck$ is an excellent predictor but tends to be slightly optimistic. Moreover, $b > ck$ is a surprisingly robust rule and returns suitable predictions even for highly inhomogeneous graphs such as scale-free networks (Albert and Barabási, 2002) but not surprisingly the deviations tend to increase (see Fig. 9b).

In principle, the pair approximation method can be applied to any kind of game (arbitrary payoff matrices) and is not limited to the notion chosen here (Ohtsuki et al., 2006). However, if the analysis is restricted to costs and benefits of cooperation in the prisoner's dilemma, the Taylor et al. (2007) have recently shown that in this special case an analysis based on inclusive fitness theory (Hamilton, 1964) naturally yields finite size corrections arising in small populations.

3.5. EQUILIBRIA

The $b > ck$ rule derived in the previous section determines the condition in the weak selection limit for evolution to favor cooperation in the prisoner's dilemma. In principle, an analogous calculation could be carried out for the snowdrift game but because cooperators and defectors easily co-exist in such interactions, fixation probabilities become less relevant because of exceedingly long fixation times (c.f. Sect. 3.2). The same holds in the spatial prisoner's dilemma if the selection w is sufficiently strong. Finally, in large populations the dynamics is dominated by deterministic drift and stochastic effects that are required for the fixation of one or the other strategic type becomes less relevant.

In all these cases it is more appropriate to consider large populations and investigate the effects of space on the equilibrium frequencies of cooperators and defectors in the prisoner's dilemma as well as the snowdrift game. In particular, we focus on square lattices where each individual has $k = 4$ neighbors. For the updating of the population we adopt a spatial analogue of the replicator dynamics (Hofbauer and Sigmund, 1998): a focal individual is randomly selected from the entire population and its fitness f_f corresponds to the average performance in interactions with all its neighbors. Second, a random neighbor of the focal individual is chosen and its fitness f_n is determined in the same way. The focal individual adopts the strategy of a better performing neighbor with a probability proportional to the fitness difference $z = f_n - f_f$ and sticks to its strategy otherwise. The transition probability can then be written as $\tau(z) = z^+/\alpha$, where $z^+ = z$ for $z > 0$ and zero otherwise and where α indicates a suitable normalization constant to ensure $\tau(z) \leq 1$. α depends on the type of interactions: for the prisoner's dilemma $\alpha = w(b + c)$ and for the snowdrift game $\alpha = wb$. Note that $\tau(z)$ is independent of the selection strength w and thus we set $w = 1$ without loss of generality. This

particular functional form of $\tau(z)$ was chosen because it recovers the replicator Eq. (7) in the limit $N, k \rightarrow \infty$. Alternatively, all individuals could be updated in synchrony to model populations with non-overlapping generations but for stochastic update rules this barely affects the equilibrium frequencies (Doebeli and Hauert, 2005).

Unfortunately it is impossible to solve the evolutionary dynamics of this system and we have to resort to simulation data. However, pair approximation provides again a welcome analytical complement and yields useful numerical estimates. Essentially by following the reasoning in the previous section and taking the different updating procedure into account, some algebra leads to the following dynamical equations:

$$\dot{p}_c = k p_c^{k-1} \sum_{i=0}^{k-1} \binom{k-1}{i} q_{cd}^i (1 - q_{cd})^{k-1-i} \sum_{j=0}^{k-1} q_{clc}^j (1 - q_{clc})^{k-1-i} \times \\ (\tau(f_d^{i+1} - f_c^j) - \tau(f_c^j - f_d^{i+1})) \quad (19a)$$

$$\dot{q}_{clc} = p_c^{k-2} \sum_{i=0}^{k-1} \binom{k-1}{i} q_{cd}^i (1 - q_{cd})^{k-1-i} \sum_{j=0}^{k-1} q_{clc}^j (1 - q_{clc})^{k-1-i} \times \\ ((2(i+1) - k q_{clc})\tau(f_d^{i+1} - f_c^j) - (2j - k q_{clc})\tau(f_c^j - f_d^{i+1})) \quad (19b)$$

where f_c^i, f_d^i denote the fitness of cooperators and defectors that have i cooperators among their k neighbors. Also recall that $q_{cd} = (1 - p_c(2 - q_{clc})) / (1 - p_c)$. Note that the sums run only up to $k - 1$ because each focal cooperator must have at least one defecting neighbor and vice versa – only in these cases changes in p_c and q_{clc} can occur. Technical details on Eq. (19) are provided in Hauert and Doebeli (2004) and Hauert and Szabó (2005).

As mentioned earlier, in the prisoner's dilemma spatial structure supports cooperation (see Fig. 10). For sufficiently low cost to benefit ratios of mutual cooperation $r = c/(b - c)$ cooperators and defectors co-exist in a dynamical equilibrium. Cooperators persist by forming compact clusters such that they are more likely to interact with other cooperators and thereby reduce exploitation by defectors. However, the clustering advantage is limited and cooperators go extinct if the benefits do not exceed the 14-fold costs. Qualitatively these results also hold for updating mechanisms based on the Moran process with marginal changes for birth-death updating and enhanced support of cooperation for death-birth updating (Hauert, 2006). Note that for the update rule inspired by the replicator dynamics $\rho_C < 1/N$ always holds, which suggests that evolution should never favor cooperation. This is no contradiction because although cooperators may never manage to reach fixation, they can nevertheless co-exist with defectors for arbitrarily long times.

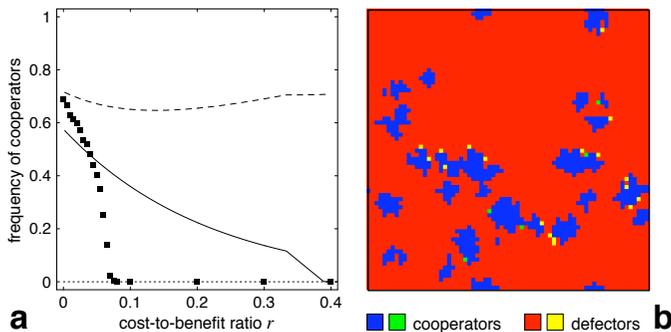


Figure 10. Spatial prisoner's dilemma on a square 100×100 lattice with von Neumann neighborhood, $k = 4$. **a** Simulated equilibrium fraction of cooperators (solid squares) as a function of the cost-to-benefit ratio of mutual cooperation $r = c/(b-c)$. For small r cooperators persist but disappear for $r > r_c \approx 0.076$. In unstructured populations, cooperators could never survive (dotted line). Pair approximation correctly predicts the increase in cooperation due to spatial structure but it greatly overestimates its effect (p_c , solid line; q_{clc} , dashed line). The consistently high values of q_{clc} indicates high degrees of clustering. **b** Snapshot of a typical lattice configuration near the extinction threshold r_c .

Near the extinction threshold small clusters of cooperators slowly meander in a sea of defectors. Occasionally two clusters collide and merge or one cluster splits into two. This resembles a branching and annihilating random walk (see Fig. 10b) and, indeed, as r approaches the extinction threshold r_c , the simulation data suggests that the system undergoes a critical phase transition that belongs to the directed percolation universality class (Szabó and Hauert, 2002b; Szabó and Hauert, 2002a).

Based on these results for the prisoner's dilemma, it was generally accepted that spatial structure promotes cooperation. However, this is not true in general for the snowdrift game. Quite on the contrary, spatial structure often inhibits the evolution of cooperation and may even eliminate cooperation altogether (see Fig. 11). In unstructured populations cooperators and defectors co-exist in a stable equilibrium because the rare type is always favored but the very same mechanism turns out to be detrimental to cooperation in spatial settings. For every individual it is always better to adopt a strategy that is different from its neighbors and this prevents the formation of larger clusters. Instead, cooperators form dendritic structures and filament-like clusters that increase interactions between cooperators and defectors (see Fig. 11b). In order to estimate the extinction threshold of cooperators we consider the threshold where the fitness of an isolated cooperator drops below the fitness of the neighboring defectors: $k(b-c) < b$, which translates to $r > (k-1)/(k+1)$. For $k = 4$ this yields a threshold of 0.6, which slightly underestimates the

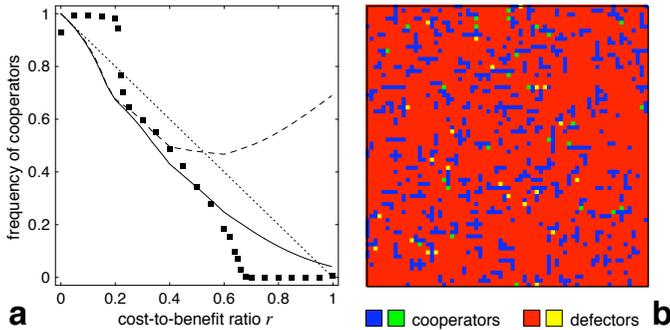


Figure 11. Spatial snowdrift game on a square 100×100 lattice with von Neumann neighborhood, $k = 4$. **a** Simulated equilibrium fraction of cooperators (solid squares) as a function of the cost-to-benefit ratio of mutual cooperation $r = c/(2b - c)$. In unstructured populations cooperators and defectors co-exist (dotted line). With the exception of small r , spatial structure inhibits cooperation and for $r > r_c \approx 0.68$ space even eliminates cooperation altogether. The overall trend to inhibit cooperation is correctly predicted by pair approximation (p_c , solid line; q_c , dashed line) but it is unable to capture the extinction of cooperators and defectors respectively. **b** Snapshot of a typical lattice configuration near the extinction threshold r_c .

extinction threshold derived from simulations with $r_c \approx 0.68$ (see Fig. 11a). For updating mechanisms based on the Moran process the detrimental effects of space are weaker and almost disappear for death-birth updating (Hauert, 2006).

4. The Origin of Cooperators and Defectors

In nature cooperation may not always be an all or nothing decision as we have assumed in the previous section by considering the evolutionary fate of two distinct strategic types, the cooperators and the defectors. Instead, in many situations it may be more appropriate to assume a continuous range of degrees of cooperation, such as time and effort expended in providing benefits to specific individuals or for the common good. In such continuous games the strategy or trait x of an individual indicates the effort or investment in cooperative interactions that can vary between zero and an upper limit x_{\max} . The associated fitness benefits and costs are specified by two functions $B(x)$ and $C(x)$. We assume that $B(x), C(x)$ are smooth and strictly increasing functions in the interval $[0, x_{\max}]$, with $B(0) = C(0) = 0$ such that zero investments into cooperation (or pure defection) incur no costs and provide no benefits.

In the traditional prisoner's dilemma, cooperators provide a benefit b to their partner at some cost c to themselves. Translating this setup to continuous strategies yields the payoff to an x -strategist interacting with a y -strategist:

$Q(x, y) = B(y) - C(x)$. The benefits are determined by the opponents strategy whereas the costs depend on the individuals own strategy. This situation applies, for example, in grooming baboons (Saunders and Hausfater, 1988; Stambach and Kummer, 1982) where one individual grooms the other for a time x and vice versa for a time y . In this case, the only way for an individual to improve its payoff is to reduce the costs and reduce the grooming time x . Consequentially, evolution selects lower investors and x readily approaches zero (Killingback and Doebeli, 2002). Cooperation disappears and defectors reign in both the traditional (c.f. Sect. 3.1) as well as the continuous prisoner's dilemma. The baboons avoid this unfortunate outcome because they ensure their partner's fidelity by taking turns in a single grooming session. However, our simple setup excludes such strategic responses and nothing prevents cooperation from disappearing. Moreover, complex behavioral patterns are only available to higher organisms and certainly do not occur in microorganisms.

In the traditional snowdrift game, cooperators also provide a benefit b to their partner but the costs c are shared among cooperators. Equivalently, we could assume that costs are fixed and benefits accumulate at a discounted rate $\gamma < 1$ (Hauert et al., 2006c), such that mutual cooperation yields $b(1 + \gamma) - c$ and the other payoffs remain unchanged (c.f. Eq. (11)). Such situations seem to apply in yeast cells that secrete enzymes in order to hydrolyze sucrose (Greig and Travisano, 2004). If only one cell produces the enzyme the resulting food resource may be critical for survival, whereas the value of additional food is discounted because the cells get saturated. Thus, if food is scarce it may be better to invest more into enzyme production and prevent starvation despite the prospects of being exploited. Conversely, if food is abundant an individual may improve its fitness by lowering enzyme production and increase reproduction. Thus, the payoff to an x -strategist interacting with a y -strategist becomes $P(x, y) = B(x + y) - C(x)$. As before, the costs are determined by the individuals own strategy but the benefits depend on the strategies of both players. In this case one could expect that strategies would evolve away from zero to an intermediate level provided that $B(x) > C(x)$ holds for small x . However, as we shall see, the continuous snowdrift game exhibits much richer evolutionary dynamics (Doebeli et al., 2004).

The continuous snowdrift game potentially applies whenever individuals produce a valuable common resource at some cost to themselves (Doebeli and Hauert, 2005), which describes a particular but abundant form of social dilemmas. Numerous examples can be found in the microbial world ranging from viruses where replication enzymes represent a common resource (Huang and Baltimore, 1977), and RNA phages producing proteins for the virus capsule (Turner and Chao, 2003), to antibiotic resistance in bacteria secreting

β -lactamase to inhibit cell wall synthesis (Neu, 1992) and the formation of fruiting bodies in *Mycococcus xanthus* (Velicer et al., 2000). Examples from higher organisms include sentinel behavior in meerkats (Clutton-Brock et al., 1999) and predator inspection in fish (Milinski, 1987; Pitcher, 1992), where the information that the inspectors obtain can be viewed as a public resource (Magurran and Higham, 1988), to cultural evolution in humans with interactions from communal enterprises to global sustainability issues (Milinski et al., 2006).

4.1. ADAPTIVE DYNAMICS

In the continuous snowdrift game the evolution of the trait x can be analyzed using the framework of adaptive dynamics (Dieckmann and Law, 1996; Geritz et al., 1998; Metz et al., 1996). Assume a homogeneous monomorphic population of x -strategists and determine whether a rare mutant with strategy y can invade. The fitness of the y strategy is simply given by $P(y, x)$ because at least as long as y -strategists are rare, interactions with other y -strategists, $P(y, y)$, can be neglected. From replicator dynamics it follows that y -strategists increase in abundance if their fitness exceeds the fitness of the resident, $P(x, x)$. Thus, the growth rate of the y -strategist is given by $f_x(y) = P(y, x) - P(x, x) = B(x + y) - C(y) - (B(2x) - C(x))$ and is called the invasion fitness because if $f_x(y) > 0$ the y mutant invades and disappears if $f_x(y) < 0$. In the limit of small mutations where y is very similar to x , it can be proven that if $f_x(y) > 0$ holds, the y -strategist not only invades but also replaces the resident population (Geritz et al., 1998). The occasional appearance of a rare mutant drives the evolutionary process but mutation rates must be small such that between subsequent invasion attempts the population has sufficient time to relax into a homogeneous state. Under these conditions, the evolution of the trait x is governed by the selection gradient $D(x) = \partial f_x / \partial y|_{y=x} = B'(2x) - C'(x)$ and the adaptive dynamics of x is described by $\dot{x} = mD(x)$ where m depends on the population size and reflects the mutational process supplying new mutant strategies. For constant population sizes m is simply a constant and is set to $m = 1$ without loss of generality. If $D(x) > 0$ mutants with $y > x$ can invade and the resident trait increases over time. Conversely, if $D(x) < 0$ mutants with $y < x$ invade and the resident trait decreases.

Equilibrium points of the adaptive dynamics, $\dot{x} = 0$, are called singular strategies x^* and are solutions to $D(x^*) = B'(2x^*) - C'(x^*) = 0$. If no such solution exists in the interval $(0, x_{\max})$, then trait values either decrease until cooperative contributions vanish ($D(x) < 0$) or keep increasing until x_{\max} is reached ($D(x) > 0$). Both situations can occur in the continuous snowdrift game (see Fig. 12): the first case is dynamically equivalent to the

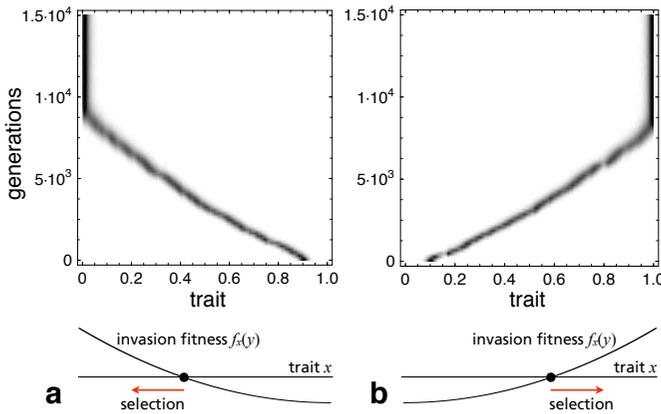


Figure 12. Dynamics in the continuous snowdrift game in the absence of singular strategies. The top row depicts simulation results for the trait distribution in the population over time, where darker shades indicate higher abundance of traits, and the bottom row provides a schematic illustration of the fitness profile in the population. The trait interval is restricted to $[0, 1]$ and the benefit and cost functions are quadratic, $B(x) = b_2x^2 + b_1x$, $C(x) = c_2x^2 + c_1x$, such that $C(x), B(x)$ are saturating and strictly increasing in $[0, 1]$. **a** The selection gradient is always negative, $D(x) < 0$, and irrespective of the initial configuration evolution keeps favoring individuals that invest less into cooperation until cooperation vanishes, just as in the continuous prisoner's dilemma. The qualitative features of the invasion fitness $f_x(y)$ do not change as x changes over time. **b** This is the exact opposite of **a**: $D(x) > 0$ always holds and the traits in the population invariably approach the maximum level of cooperation. Parameters: $b_2 = -1.5, b_1 = 7, c_2 = -1$ and **a** $c_1 = 8$; **b** $c_1 = 2$.

continuous prisoner's dilemma whereas in the second case full cooperation is established and is sometimes termed by-product mutualism because increasing cooperation yields a net benefit to the actor and benefits to others occur only as a side effect (Connor, 1996; Dugatkin, 1996; Milinski, 1996). The dynamics becomes more interesting if x^* exists. The singular strategy is convergent stable and hence an attractor of the evolutionary dynamics if $dD(x)/dx|_{x=x^*} < 0$ and is an evolutionary repeller if this inequality is reversed, i.e. the trait x evolves away from x^* (see Fig. 13c). If x^* is an attractor, the traits in the population converge to x^* but the subsequent evolutionary fate of the population depends on whether x^* is also evolutionarily stable, i.e. whether x^* denotes a maximum or minimum of the invasion fitness $f_x(y)$. If $\partial^2 f_x(y)/\partial y^2|_{y=x^*} = 2B''(2x^*) - C''(x^*) < 0$ then x^* represents a fitness maximum and thus represents an evolutionary end state where every individual provides equal intermediate cooperative contributions and corresponds to the original expectation (see Fig. 13b). If, however, $2B''(2x^*) - C''(x^*) > 0$ then a population of x^* -strategists is at a fitness minimum and mutants with either

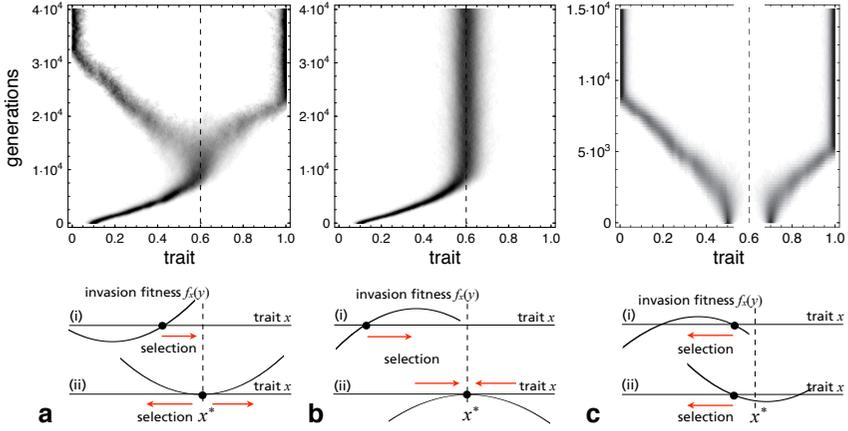


Figure 13. Dynamics in the continuous snowdrift game in presence of a unique singular strategy x^* . The top row depicts simulation results for the trait distribution in the population over time, where darker shades indicate higher abundance of traits, and the bottom row provides a schematic illustration of the fitness profile in the population. x^* is marked by a vertical dashed line. As in Fig. 12, the trait interval is $[0, 1]$ and the benefit and cost functions are quadratic, saturating and strictly increasing in $[0, 1]$. **a** *evolutionary branching* – the singular strategy is convergent stable and the trait distribution approaches x^* but it is not evolutionarily stable and the population branches into two distinct phenotypic clusters. Evolution selects individuals with higher fitness (bottom panel (i)) but this also changes the profile of the invasion fitness $f_x(y)$ such that the fitness minimum catches up at x^* (bottom panel (ii)) such that mutants with both higher and lower y can invade. **b** *evolutionary stability* – the singular strategy is not only convergent stable but also evolutionarily stable such that the trait distribution approaches x^* and remains there. As the population converges to x^* (bottom panel (i)) the profile of $f_x(y)$ changes and at x^* the trait catches up with the maximum of $f_x(y)$ (bottom panel (ii)) and no mutants are able to invade. **c** *evolutionary repeller* – the singular strategy is an evolutionary repeller such that the traits evolve away from x^* . Two separate simulation runs are shown: when starting below x^* cooperation disappears but if initial cooperative contributions are sufficiently high they keep increasing until the maximum is reached. In this case it is irrelevant whether x^* is evolutionarily stable (bottom panel (i)) or an evolutionary branching point (bottom panel (ii)) because evolution never reaches x^* and would require careful preparations of the initial configuration. Parameters: **a** $b_2 = -1.4, b_1 = 6, c_2 = -1.6, c_1 = 4.56$; **b** $b_2 = -1.5, b_1 = 7, c_2 = -1, c_1 = 4.6$; **c** $b_2 = -0.5, b_1 = 3.4, c_2 = -1.5, c_1 = 4$.

higher or lower traits y can invade. In this case the population undergoes evolutionary branching and spontaneously splits into two distinct phenotypic clusters of high and low investing individuals (see Fig. 13a).

For quadratic benefit and cost functions $B(x) = b_2x^2 + b_1x$, $C(x) = c_2x^2 + c_1x$ with suitable parameters, such that both benefits and costs are strictly increasing over the trait interval $[0, x_{\max}]$, all dynamical scenarios occur (see Figs. 12, 13) and can be fully analyzed (Doebeli et al., 2004). In this case,

the singular strategy x^* is unique (if it exists) and is given by $x^* = (c_1 - b_1)/(4b_2 - 2c_2)$. It is convergent stable if $2b_2 - c_2 < 0$ and evolutionarily stable if $b_2 - c_2 < 0$. The existence of x^* requires either (i) $4b_2 - 2c_2 > c_1 - b_1 > 0$ or (ii) $4b_2 - 2c_2 < c_1 - b_1 < 0$. In the first case x^* is always a repeller and whether it is evolutionarily stable is irrelevant because the singular strategy is never reached from generic initial conditions. In the second case x^* is always convergent stable and if, in addition, $b_2 - c_2 < 0$ holds it is also evolutionarily stable. Finally, if $2b_2 < c_2 < b_2 < 0$ then x^* is an evolutionary branching point. In order to analyze the dynamics after branching has occurred, the invasion fitness needs to be derived for a third strategic type z attempting to invade a resident population where the two co-existing strategies $x > x^* > y$ are in equilibrium. This leads to two dynamical equations one for each branch x and y (Doebeli et al., 2004). In the case of quadratic cost and benefit functions the traits keep diverging until they reach the boundaries of the trait interval because mutants with either $z > x$ or $z < y$ can invade.

Note that the phenotypic diversification occurs spontaneously and in populations with randomly interacting individuals and thus does not rely on any kind of assortment in terms of mating preferences or spatial segregation. The evolutionary end state consist of a population with pure defectors and pure cooperators that, in fact, engage in traditional snowdrift game interactions (c.f. Sect. 3.2). Thus, the continuous snowdrift game suggest an evolutionary pathway for social diversification and for the origin of cooperators and defectors.

4.2. TWO TRAGEDIES

The conflict of interest in social dilemmas is equivalently captured by the *Tragedy of the Commons* (Hardin, 1968), which states that public resources are bound to be overexploited. Especially in the context of humans, this comes at no surprise – Aristotle (384-322 BC) already drew the same conclusion: "That which is common to the greatest number has the least care bestowed upon it." The spontaneous diversification into co-existing high and low investors in the continuous snowdrift game may equally apply in communal enterprises in humans and generate, in addition, the *Tragedy of the Commune* (Doebeli et al., 2004), which states that evolution may not favor egalitarian contributions to the common good but instead promote highly asymmetric involvements. But, large differences in cooperative contributions bear a formidable potential for escalating conflicts based on the accepted notion of fairness.

5. Conclusions

This brief review discusses different approaches to model evolutionary dynamics and address the problem of cooperation. Despite the simple principles underlying evolutionary models, they often exhibit rich, complex and sometimes even chaotic but always truly fascinating dynamics. Because of this, it can be quite challenging to develop an intuitive understanding of the dynamical features. In order to achieve this intuition it is very helpful to study characteristics of spatio-temporal patterns or to follow the evolutionary change of strategy distributions in a population. This is achieved by the *VirtualLabs* (Hauert, 2007), which complement the research results presented in this review. The growing collection of interactive on-line tutorials comprises at least one tutorial for each section covered in this review. Based on Java applets, most of the results reported here can be easily reproduced and further explorations are encouraged by the possibility to change various settings. But it is also possible to simply watch and enjoy the hypnotizing beauty of evolutionary kaleidoscopes – find out what they are by visiting the *VirtualLabs*.

Acknowledgements

Support by the John Templeton Foundation and the NSF/NIH joint program in mathematical biology (NIH grant R01GM078986) is gratefully acknowledged. The Program for Evolutionary Dynamics (PED) at Harvard University is sponsored by Jeffrey Epstein.

References

- Albert, R. and Barabási, A.-L. (2002) Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74**, 47–97.
- Barabási, A.-L. and Albert, R. (1999) Emergence of Scaling in Random Networks, *Science* **286**, 509–512.
- Bollobás, B. (1995) *Random Graphs*, New York, Academic Press.
- Clutton-Brock, T. H., M. J. O’Riain, P. N. M. B., Gaynor, D., Kansky, R., Griffin, A. S., and Manser, M. (1999) Selfish Sentinels in Cooperative Mammals, *Science* **284**, 1640–1644.
- Clutton-Brock, T. H. and Parker, G. A. (1995) Punishment in animal societies, *Nature* **373**, 209–216.
- Connor, R. C. (1996) Partner preferences in by-product mutualisms and the case of predator inspection in fish, *Anim. Behav.* **51**, 451–454.
- Dawes, R. M. (1980) Social Dilemmas, *Ann. Rev. Psychol.* **31**, 169–193.
- Dieckmann, U. and Law, R. (1996) The dynamical theory of coevolution: a derivation from stochastic ecological processes, *J. Math. Biol.* **34**, 579–612.
- Doebeli, M. and Hauert, C. (2005) Models of cooperation based on the Prisoner’s Dilemma and the Snowdrift game, *Ecol. Lett.* **8**, 748–766.

- Doebeli, M., Hauert, C., and Killingback, T. (2004) The evolutionary origin of cooperators and defectors, *Science* **306**, 859–862.
- Dugatkin, L. A. (1996) Tit for Tat, by-product mutualism and predator inspection: a reply to Connor, *Anim. Behav.* **51**, 455–457.
- Flood, M. (1958) Some Experimental Games, *Management Science* **5**, 5–26.
- Geritz, S. A. H., Kisdi, E., Meszéna, G., and Metz, J. A. J. (1998) Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree, *Evol. Ecol.* **12**, 35–57.
- Greig, D. and Travisano, M. (2004) The Prisoner’s Dilemma and polymorphism in yeast SUC genes, *Biol. Lett.* **271**, S25 – S26.
- Hamilton, W. D. (1964) The genetical evolution of social behaviour I, *J. theor. Biol.* **7**, 1–16.
- Hamilton, W. D. (1971) The Geometry of the Selfish Herd, *J. theor. Biol.* **31**, 295–311.
- Hardin, G. (1968) The tragedy of the commons, *Science* **162**, 1243–1248.
- Hauert, C. (2006) Spatial effects in social dilemmas, *J. theor. Biol.* **240**, 627–636.
- Hauert, C. (2007) VirtualLabs: Interactive Tutorials on Evolutionary Game Theory, <http://www.univie.ac.at/virtuallabs>.
- Hauert, C., De Monte, S., Hofbauer, J., and Sigmund, K. (2002) Volunteering as red queen mechanism for cooperation in public goods games, *Science* **296**, 1129–1132.
- Hauert, C. and Doebeli, M. (2004) Spatial structure often inhibits the evolution of cooperation in the Snowdrift game, *Nature* **428**, 643–646.
- Hauert, C., Holmes, M., and Doebeli, M. (2006)a Evolutionary games and population dynamics: maintenance of cooperation in public goods games, *Proc. R. Soc. Lond. B* **273**, 2565–2570.
- Hauert, C., Holmes, M., and Doebeli, M. (2006)b Evolutionary games and population dynamics: maintenance of cooperation in public goods games, *Proc. R. Soc. Lond. B* **273**, 3131–3132.
- Hauert, C., Michor, F., Nowak, M., and Doebeli, M. (2006)c Synergy and discounting of cooperation in social dilemmas, *J. theor. Biol.* **239**, 195–202.
- Hauert, C. and Szabó, G. (2005) Game theory and physics, *Am. J. Phys.* **73**, 405–414.
- Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., and Sigmund, K. (2007) Via freedom to coercion: the emergence of costly punishment, *Science* **316**, 1905–1907.
- Hofbauer, J. and Sigmund, K. (1998) *Evolutionary Games and Population Dynamics*, Cambridge, Cambridge University Press.
- Huang, A. S. and Baltimore, D. (1977) *Comprehensive Virology*, Vol. 10, Chapt. Defective Interfering Animal Viruses, pp. 73–116, New York, , Plenum Press.
- Karlin, S. and Taylor, H. M. (1975) *First Course in Stochastic Processes*, London, Academic Press, 2nd edition.
- Killingback, T. and Doebeli, M. (2002) The Continuous Prisoner’s Dilemma and the Evolution of Cooperation through Reciprocal Altruism with Variable Investment, *Am. Nat.* **160**, 421–438.
- Koonin, E. V., Wolf, Y. I., and Karev, G. P. (2006) *Power Laws, Scale-Free Networks and Genome Biology*, Molecular Biology Intelligence Unit, Springer.
- Lieberman, E., Hauert, C., and Nowak, M. A. (2005) Evolutionary dynamics on graphs, *Nature* **455**, 312–316.
- Magurran, A. E. and Higham, A. (1988) Information transfer across fish shoals under predator threat, *Ethology* **78**, 153–158.
- Maruyama, T. (1970) Effective number of alleles in a subdivided population, *Theor. Pop. Biol.* **1**, 273–306.

- Matsuda, H., Ogita, N., Sasaki, A., and Sato, K. (1992) Statistical Mechanics of Populations, *Prog. theor. Phys.* **88**, 1035–1049.
- Maynard Smith, J. (1982) *Evolution and the theory of games*, Cambridge, Cambridge University Press.
- Maynard Smith, J. and Price, G. (1973) The logic of animal conflict, *Nature* **246**, 15–18.
- Maynard Smith, J. and Szathmáry, E. (1995) *The major transitions in evolution*, Oxford, W. H. Freeman & Co.
- Metz, J. A. J., Geritz, S. A. H., Meszina, G., Jacobs, F. J. A., and van Heerwaarden, J. S. (1996) Adaptive dynamics: a geometrical study of the consequences of nearly faithful replication. In S. J. van Strien and S. M. Verduyn Lunel (eds.), *Stochastic and Spatial Structures of Dynamical Systems*, Amsterdam, North Holland, pp. 183–231.
- 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., Semmann, D., Krambeck, H.-J., and Marotzke, M. (2006) Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments, *Proc. Natl. Acad. Sci. USA* **103**, 3994–3998.
- Moran, P. A. P. (1962) *The Statistical Processes of Evolutionary Theory*, Oxford, UK, Clarendon.
- Nash, J. (1951) Non-cooperative games, *Annals of Mathematics* **54**, 286–299.
- Neu, H. C. (1992) The Crisis in Antibiotic Resistance, *Science* **257**, 1064–1073.
- Nowak, M. A. (2006) *Evolutionary Dynamics*, Cambridge MA, Harvard University Press.
- Nowak, M. A. and May, R. M. (1992) Evolutionary Games and Spatial Chaos, *Nature* **359**, 826–829.
- Nowak, M. A., Michor, F., and Iwasa, Y. (2003) The linear process of somatic evolution, *Proc. Natl. Acad. Sci. USA* **100**, 14966–14969.
- Nowak, M. A., Sasaki, A., Taylor, C., and Fudenberg, D. (2004) Emergence of cooperation and evolutionary stability in finite populations, *Nature* **428**, 646–650.
- Nowak, M. A. and Sigmund, K. (1998) Evolution of indirect reciprocity by image scoring, *Nature* **393**, 573–577.
- Ohtsuki, H., Hauert, C., Lieberman, E., and Nowak, M. A. (2006) A simple rule for the evolution of cooperation on graphs and social networks, *Nature* **441**.
- Pitcher, T. (1992) Who Dares, Wins - the Function and Evolution of Predator Inspection Behavior in Shoaling Fish, *Neth. J. Zool.* **42**, 371–391.
- Poundstone, W. (1992) *Prisoner's Dilemma*, New York, Doubleday.
- Rainey, P. B. and Rainey, K. (2003) Evolution of cooperation and conflict in experimental bacterial populations, *Nature* **425**, 72–74.
- Saunders, C. D. and Hausfater, G. (1988) The Functional Significance of Baboon Grooming Behavior, *Annals N.Y. Acad. Sci.* **525**, 430–432.
- Slatkin, M. (1981) Fixation probabilities and fixation times in a subdivided population, *Evolution* **35**, 477–488.
- Stammach, E. and Kummer, H. (1982) Individual contributions to a dyadic interaction: An analysis of baboon grooming, *Anim. Behav.* **30**, 964–971.
- Sugden, R. (1986) *The Economics of Rights, Co-operation and Welfare*, Oxford and New York, Blackwell.
- Szabó, G. and Hauert, C. (2002) a Evolutionary prisoner's dilemma with optional participation, *Phys. Rev. E* **66**, 062903.

- Szabó, G. and Hauert, C. (2002)b Phase transitions and volunteering in spatial public goods games, *Phys. Rev. Lett.* **89**, 118101.
- Taylor, C., Iwasa, Y., and Nowak, M. A. (2006) A symmetry of fixation times in evolutionary dynamics, *J. theor. Biol.* **243**, 245–251.
- Taylor, P. D., Day, T., and Wild, G. (2007) Evolution of cooperation in a finite homogeneous graph, *Nature* p. (in print).
- Traulsen, A., Claussen, J. C., and Hauert, C. (2005) Coevolutionary Dynamics: From Finite to Infinite Populations, *Phys. Rev. Lett.* **95**, 238701.
- Trivers, R. L. (1971) The evolution of reciprocal altruism, *Q. Rev. Biol.* **46**, 35–57.
- Turner, P. E. and Chao, L. (1999) Prisoner's Dilemma in an RNA virus, *Nature* **398**, 441–443.
- Turner, P. E. and Chao, L. (2003) Escape from Prisoner's Dilemma in RNA Phage $\Phi 6$, *Am. Nat.* **161**, 497–505.
- van Baalen, M. and Rand, D. A. (1998) The unit of selection in viscous populations and the evolution of altruism, *J. theor. Biol.* **193**, 631–648.
- Velicer, G. J., Kroos, L., and Lenski, R. E. (2000) Developmental cheating in the social bacterium *Myxococcus xanthus*, *Nature* **404**, 598–601.
- Vogelstein, B. and Kinzler, K. W. (1998) *The Genetic Basis of Human Cancer*, Toronto, McGraw-Hill.
- Watts, D. J. and Strogatz, S. H. (1998) Collective dynamics of 'small world' networks, *Nature* **393**, 440–442.