Pattern Formation, Social Forces, and Diffusion Instability in Games with Success-Driven Motion

Dirk Helbing

ETH Zurich, UNO D11, Universitätstr. 41, 8092 Zurich, Switzerland

Received: date / Revised version: date

Abstract. It is shown that success-driven motion can support the survival or spreading of cooperation through pattern formation, even where cooperation is predicted to die out according to the replicator equation, which is often used in evolutionary game theory to study the spreading and disappearance of strategies. Success-driven motion is formulated here as a function of the game-theoretical payoffs. It can change the outcome and dynamics of spatial games dramatically, in particular as it causes attractive or repulsive interaction forces. These forces act when the spatial distributions of strategies are inhomogeneous. However, even when starting with homogeneous initial conditions, small perturbations can trigger large inhomogeneities by a pattern-formation instability, when certain conditions are fulfilled. Here, these instability conditions are studied for the prisoner’s dilemma and the snowdrift game. A positive side effect of the pattern formation instability is that it can support the formation of clusters and, thereby, significantly enhance the level of cooperation in the prisoner’s dilemma and other games. Therefore, success-driven motion may be used to supplement other mechanisms supporting cooperation, like reputation or punishment. Furthermore, diffusion can sometimes trigger pattern-formation in social, economic, and biological systems, by driving the system into the unstable regime.

PACS. 02.50.Le Decision theory and game theory – 87.23.Ge Dynamics of social systems 82.40.CkPattern formation in reactions with diffusion, flow and heat transfer – 87.23.Cc Population dynamics and ecological pattern formation

1 Introduction

Game theory is a well-established theory of individual strategic interactions with applications in sociology, economics, and biology [1–9], and with many publications even in physics (see Ref. [10] for an overview). It distinguishes different behaviors, so-called strategies $i$, and expresses the interactions of individuals in terms of payoffs $P_{ij}$. The value $P_{ij}$ quantifies the result of an interaction between strategies $i$ and $j$ for the individual pursuing strategy $i$. The more favorable the outcome of the interaction is, the higher the payoff $P_{ij}$.

There are many different games, depending on the structure of the payoffs, the social interaction network, the number of interaction partners, the frequency of interaction, and so on [4,6]. Theoretical predictions for the selection of strategies mostly assume a rational choice approach, i.e. a payoff maximization by the individuals, although experimental studies [11–13] support conditional cooperativity [14] and show that moral sentiments [15] can support cooperation. Some models also take into account learning (see, e.g. [16] and references therein), where it is common to assume that more successful behaviors are imitated (copied). Based on a suitable specification of the imitation rules, it can be shown [17–19] that the resulting dynamics can be described by game-dynamical equations [20–23], which agree with replicator equations for the fitness-dependent reproduction of individuals in biology [24–26].

Another field where the quantification of human behavior in terms of mathematical models has been extremely successful concerns the dynamics of pedestrians [27], crowds [28], and traffic [29]. The related studies have led to fundamental insights into observed self-organization phenomena such as stop-and-go waves [29] or lanes of uniform walking direction [27]. In the meantime, there are many empirical [30] and experimental results [31], which made it possible to come up with well calibrated models of human motion [33,34].

Therefore, it would be interesting to know what happens if game theoretical models are combined with models of driven motion. Would we also observe self-organization phenomena in space and time? This is the main question addressed in this paper, which is structured as follows: Section 2 introduces the game-dynamical equations without spatial interactions and specifies the payoffs and model parameters for the prisoner’s dilemma and the snowdrift game. While this section serves the illustration of the problem of emergent cooperation along the lines previously studied in the literature, we will extend the model...
in Sec. 3 in order to take spatial interactions into account. In particular, we will introduce success-driven motion and diffusion. The analysis of the model in Sec. 3.1 reveals that the dynamics in games with success-driven motion can be described by (social) interaction forces. Afterwards, in Sec. 3.2, we will perform a linear stability analysis and show that games with success-driven motion tend to be unstable and to produce emergent patterns, such as clusters.

Under keywords such as “assortment” and “network reciprocity”, it has been discussed that the clustering of cooperators can amplify cooperation, in particular in the prisoner’s dilemma [35–38]. Therefore, the pattern formation instability is of prime importance to understand the emergence of cooperation between individuals. In Sec. 4, we will study the instability conditions for the prisoner’s dilemma and the snowdrift game. Moreover, we will see that games with success-driven motion and diffusion are expected to show pattern formation, where a homogeneous distribution of strategies would be stable without the presence of diffusion. It is quite surprising that sources of noise like diffusion can support the self-organization in systems, which can be described by game-dynamical equations with success-driven motion. This includes social, economic, and biological systems.

2 The Prisoner’s Dilemma without Spatial Interactions

In order to grasp the major impact of success-driven motion on the dynamics of spatial games (see Sec. 3), it is useful to investigate first the game-dynamical equations without spatial interactions. For this, we represent the proportion of individuals using a strategy \( i \) at time \( t \) by \( p(i, t) \). While the discussion can be extended to any number of strategies, we will focus on the case of two strategies only for the sake of analytical tractability. Here, \( i = 1 \) shall correspond to the cooperative strategy, \( i = 2 \) to defection. According to the definition of probabilities, we have \( 0 \leq p(i, t) \leq 1 \) for \( i \in \{1, 2\} \) and the normalization condition

\[
p(1, t) + p(2, t) = 1 .
\]

(1)

Let \( P_{ij} \) be the payoff, if strategy \( i \) meets strategy \( j \). Then, the expected payoff for someone applying strategy \( i \) is

\[
E(i, t) = \sum_{j=1}^{2} P_{ij} p(j, t) ,
\]

(2)

as \( p(j, t) \) represents the proportion of strategy \( j \), with which the payoffs \( P_{ij} \) must be weighted. The average payoff in the population of individuals is

\[
\bar{E}(t) = \sum_{l=1}^{2} E(l, t)p(l, t) = \sum_{l=1}^{2} \sum_{j=1}^{2} p(l, t)P_{lj} p(j, t) .
\]

(3)

In the game-dynamical equations, the temporal increase \( dp(i, t)/dt \) of the proportion of individuals using strategy \( i \) is proportional to the number of individuals pursuing strategy \( i \) who may imitate, i.e. basically to \( p(i, t) \). The proportionality factor, the growth rate \( \lambda(i, t) \), is given by the difference between the expected payoff \( E(i, t) \) and the average payoff \( \bar{E}(t) \):

\[
\frac{dp(i, t)}{dt} = \lambda(i, t)p(i, t) = [E(i, t) - \bar{E}(t)]p(i, t)
\]

\[
= \left( \sum_{j=1}^{2} P_{ij}p(j, t) - \sum_{l=1}^{2} \sum_{j=1}^{2} p(l, t)P_{lj} p(j, t) \right) p(i, t) .
\]

(4)

The equations (4) are known as replicator equations. They were developed in evolutionary biology to describe the spreading of “fitter” individuals through their higher reproduction success [24–26]. However, the replicator equations were also used in game theory, where they are called “game-dynamical equations” [20,21]. For a long time, it was not clear whether or why the application of these equations to the frequency \( p(i, t) \) of behavioral strategies was justified, but it has been shown that the equations can be derived from Boltzmann-like equations for imitative pair interactions of individuals, if “proportional imitation” or similar imitation rules are assumed [17–19].

It can be shown that

\[
\sum_{i=1}^{2} \frac{dp(i, t)}{dt} = 0 ,
\]

(5)

so that the normalization condition (1) is fulfilled at all times \( t \), if it is fulfilled at \( t = 0 \). Moreover, the equation \( dp(i, t)/dt = \lambda p(i, t) \) implies \( p(i, t) \geq 0 \) at all times \( t \), if \( p(i, 0) \geq 0 \) for all strategies \( i \).

We may now insert the payoffs of the prisoner’s dilemma, i.e. \( P_{11} = R, P_{12} = S, P_{21} = T, \) and \( P_{22} = P \) with

\[
T > R > P > S .
\]

(6)

Additionally, one often requires \( 2R > S+T \). The “reward” \( R \) is the payoff for mutual cooperation and the “punishment” \( P \) the payoff for mutual defection, while \( T \) is the “temptation” of unilateral defection, and a cheated cooperator receives the sucker’s payoff \( S \). While we have \( P > S \) in the prisoner’s dilemma, the snowdrift game (also known as chicken or hawk-dove game) is characterized by \( S > P \), i.e. it is defined by

\[
T > R > S > P .
\]

(7)

Inserting this into Eq. (4), we get

\[
\frac{dp(1, t)}{dt} = \left\{ Rp(1, t) + Sp(2, t) - R[p(1, t)]^2 \right. \\
- (S + T)p(1, t)p(2, t) - P[p(2, t)]^2 \right\} p(1, t) .
\]

(8)

and considering Eq. (1), i.e. \( p(2, t) = 1 - p(1, t) \), we find

\[
\frac{dp(1, t)}{dt} = \left\{ Rp(1, t) + S[1 - p(1, t)] - R[p(1, t)]^2 \right. \\
- (S + T)p(1, t)(1 - p(1, t)) - P[p(2, t)]^2 \right\} p(1, t) .
\]
\[ - (S + T)p(1,t)[1 - p(1,t)] \\
\frac{dp(1,t)}{dt} = [1 - p(1,t)] [ - A + Bp(1,t)] p(1,t) \\
+ (P + R - S - T)p(1,t) p(1,t). \quad (9) \]

This game-dynamical equation is a mean-value equation, which assumes a factorization of joint probabilities, i.e. it neglects correlations [19]. Nevertheless, the following analysis is suited to provide insights into the dynamics of the prisoner’s dilemma. Introducing the abbreviations

\[ A = P - S \quad \text{and} \quad B = P + R - S - T, \quad (10) \]

Eq. (9) can be further simplified, and we get

\[ \frac{dp(1,t)}{dt} = [1 - p(1,t)] [ - A + Bp(1,t)] p(1,t). \quad (11) \]

Obviously, shifting all payoffs \( P_{ij} \) by a constant value \( c \) does not change Eq. (11), in contrast to the case involving spatial interactions discussed later.

The stationary solutions of Equation (11) can be determined by setting \( \frac{dp(i,t)}{dt} = 0 \). We find three stationary solutions \( p(1,t) = p_k(1) \), namely

\[ p_1(1) = 0, \quad p_2(1) = 1 \quad \text{and} \quad p_3(1) = \frac{A}{B}. \quad (12) \]

Note that \( A - B = T - R > 0 \) and \( A > 0 \) in the prisoner’s dilemma, while \( A < 0 \) in the snowdrift game. Hence, in the prisoner’s dilemma we have \( p_3(1) = A/B > 1 \) (if \( B > 0 \)) or \( p_3(1) = A/B < 0 \) (if \( B < 0 \)), so that \( p_3(k) \) does not fall into the required range between 0 and 1. Therefore, \( p_3(1) \) cannot be a stationary solution of the prisoner’s dilemma. The situation is different for the snowdrift game, where \( p_3(1) = A/B < 1 \) due to \( B < A \). Note that one may add a mutation term such as

\[ Wp(2,t) - Wp(1,t) = W[1 - p(1,t)] - Wp(1,t) = 2W \left( \frac{1}{2} - p(1,t) \right) \quad (13) \]

to the right-hand side of the game-dynamical equations (11), where \( W \) denotes the mutation rate [19]. This implementation reflects spontaneous strategy mutations and modifies the stationary solutions. Specifically, \( p_1(1) \) will assume a finite value, which is close to zero for small values of \( W \) and converges to 1/2 in the limit \( W \to \infty \). Therefore, mutations increase the level \( p_1(1) = 0 \) of cooperation from zero to a finite value.

It will now be important to find out into which of these stationary solutions the system will evolve, i.e. which one is stable. For this, one typically performs a linear stability analysis as follows: First, we define deviations \( \delta p(1,t) = p(1,t) - p_k(1) \) from the stationary state \( p_k(1) \). This may be inserted into Eq. (11) to give

\[ \frac{d\delta p(1,t)}{dt} = [1 - p_k(1) - \delta p(1,t)] \]
\[ \times [-A + Bp_k(1) + B \delta p(1,t)] \]
\[ \times [p_k(1) + \delta p(1,t)] \]
\[ = \left\{ [1 - p_k(1)]p_k(1) + [1 - 2p_k(1)] \delta p(1,t) - [\delta p(1,t)]^2 \right\} \]
\[ \times [-A + Bp_k(1) + B \delta p(1,t)]. \quad (14) \]

If we concentrate on sufficiently small deviations \( \delta p(1,t) \), terms containing factors \( [\delta p(1,t)]^m \) with an integer exponent \( m > 1 \) can be considered much smaller than terms containing a factor \( \delta p(1,t) \). Therefore, we may linearize the above equations by dropping higher-order terms proportional to \( [\delta p(1,t)]^m \) with \( m > 1 \). This gives

\[ \frac{d\delta p(1,t)}{dt} = \left\{ [1 - p_k(1)]p_k(1) + [1 - 2p_k(1)] \delta p(1,t) \right\} \]
\[ \times [-A + Bp_k(1) + B \delta p(1,t)] \]
\[ = [1 - 2p_k(1)] \delta p(1,t)[-A + Bp_k(1)] \]
\[ + [1 - p_k(1)]p_k(1)B \delta p(1,t), \quad (15) \]

as \( [1 - p_k(1)]p_k(1)[-A + Bp_k(1)] = 0 \) for all stationary solutions \( p_k(1) \). With the abbreviation

\[ \lambda_k = [1 - 2p_k(1)][-A + Bp_k(1)] + [1 - p_k(1)]p_k(1)B, \quad (16) \]

we can write

\[ \frac{d\delta p(1,t)}{dt} = \lambda_k \delta p(1,t). \quad (17) \]

If \( \lambda_k < 0 \), the deviation \( \delta p(1,t) \) will exponentially decay with time, i.e. the solution will converge to the stationary solution \( p_k(1) \), which implies its stability. If \( \lambda_k > 0 \), however, the deviation will grow in time, and the stationary solution \( p_k(1) \) is unstable. For \( p_1(1) = 0 \), \( p_2(1) = 1 \), and \( p_3(1) = A/B \), we can easily find

\[ \lambda_1 = -A, \quad \lambda_2 = A - B, \quad \text{and} \quad \lambda_3 = A \left( 1 - \frac{A}{B} \right), \quad (18) \]

respectively. Due to \( A - B = T - R > 0 \), the solution \( p_2(1) = 1 \) corresponding to 100% cooperators is unstable with respect to perturbations, i.e. it will not persist. Moreover, the solution \( p_1(1) = 0 \) will be stable in the prisoner’s dilemma because of \( A = P - S > 0 \). In case of no mutation \( (W = 0) \), this corresponds to 0% cooperators and 100% defectors, which agrees with the expected result for the one-shot prisoner’s dilemma (if individuals decide according to rational choice theory). In the snowdrift game, however, the stationary solution \( p_1(1) = 0 \) is unstable due to \( A = P - S < 0 \), while the additional stationary solution \( p_3(1) = A/B < 1 \) is stable. Hence, in the snowdrift game we expect the establishment of a fraction \( A/B \) of cooperators.

In summary, for the prisoner’s dilemma, there is no evolutionarily stable solution with a finite percentage of cooperators, if we do not consider spontaneous strategy mutations (and neglect the effect of spatial correlations through the applied factorization assumption). According
to the above, cooperation in the PD is essentially expected to disappear. In the following, we will show that this conclusion does not hold for populations where cooperators and defectors are allowed to form patterns in space, which corresponds to an inhomogeneous distribution of cooperators and defectors. Moreover, we will demonstrate that random motion (“diffusion” in space) stabilizes the stationary solution \( p_1(1) \) with 0% cooperators (or a small percentage of cooperators in the presence of spontaneous strategy mutations), while success-driven motion can destabilize this solution, which gives rise to spatial pattern formation in the population. Together with the well-known fact that a clustering of cooperators can promote cooperation [36–38,35], pattern formation can amplify the level of cooperation, as was shown in Ref. [39].

### 3 Taking into Account Success-Driven Motion and Diffusion in Space

We now assume that individuals are distributed over different locations \( x \in [0, L] \) of a one-dimensional space. A generalization to multi-dimensional spaces is easily possible. Without loss of generality, we assume \( L = 1 \) (which can always be reached by a scaling of \( x \) by \( L \)). In the following, the proportion of individuals using strategy \( i \) at time \( t \) and at a location between \( x \) and \( x + dx \) is represented by \( p(i,x,t)dx \) with \( p(i,x,t) \geq 0 \). Due to the spatial degrees of freedom, the proportion of defectors is not immediately given by the the proportion of cooperators anymore, and the normalization condition (1) is replaced by the less restrictive equation

\[
\sum_{i=1}^{2} \int_{0}^{L} dx p(i,x,t) = 1. \tag{19}
\]

If \( \rho(i,x,t) = p(i,x,t)N/L \) represents the density of individuals pursuing strategy \( i \) at location \( x \) and time \( t \), we can also transfer this into the form

\[
\sum_{i=1}^{2} \int_{0}^{L} dx \rho(i,x,t) = \frac{N}{L} = \rho, \tag{20}
\]

where \( N \) is the total number of individuals in the system and \( \rho \) their average density.

Note that one may also consider to treat unoccupied space formally like a third strategy \( i = 0 \). In this case, however, the probabilities \( p(i,x,t) \) add up to one in all locations, which means

\[
p(0,x,t) = 1 - p(1,x,t) - p(2,x,t) \tag{21}
\]

and

\[
\frac{\partial p(0,x,t)}{\partial t} = \frac{-\partial p(1,x,t)}{\partial t} - \frac{\partial p(2,x,t)}{\partial t}. \tag{22}
\]

Therefore, \( p(0,x,t) \) can be eliminated from the system of equations, because unoccupied space does not interact with strategies 1 and 2. As a consequence, the dynamics in spatial games with success-driven motion is different from spatial games considering volunteering [40].

Let us now extend the game-dynamical equations according to

\[
\frac{\partial p(i,x,t)}{\partial t} = \left[ \sum_{j=1}^{2} \frac{p_{ij}p(j,x,t)}{p(i,x,t)} - \sum_{j=1}^{2} \int_{0}^{L} p(j,x,t)p(i,j,x,t) \right] \rho(i,x,t) - \rho \frac{\partial E(i,x,t)}{\partial x} + D \rho \frac{\partial^2 p(i,x,t)}{\partial x^2} \tag{23}
\]

with the local expected success

\[
E(i,x,t) = \sum_{j=1}^{2} P_{ij}p(j,x,t), \tag{24}
\]

compare Eq. (2). \( \partial p(i,x,t)/\partial t \) represents the (partial) time derivative, while \( D \) are diffusion constants. Assuming an additional smoothing term \( D \rho \frac{\partial^2 p(i,x,t)}{\partial x^2} \) with a small constant \( D > 0 \) on the right-hand side of Eq. (23) makes the numerical solution of this model well-behaved. In order to take into account capacity constraints (saturation effects), one could introduce a prefactor

\[
C(x,t) = 1 - \sum_{i=1}^{2} \frac{p(i,x,t)N}{\rho_{\max}L} \geq 0, \tag{25}
\]

where \( \rho_{\max} \geq N/L > 0 \) represents the maximum density of individuals. In the following, however, we will focus on the case \( C = 1 \), which allows a local accumulation of individuals.

Equation (23) can be generalized to multi-dimensional spaces by replacing the spatial derivative by the divergence operator. However, our main results can already be demonstrated for the simpler, one-dimensional case discussed here. The first term in square brackets on the right-hand side of Eq. (23) assumes that locally, an imitation of more successful strategies occurs. The second term models success-driven motion [41,42], and the last term represents diffusion. Note that, for simplicity, the specifications of imitation and success-driven motion are based on local interactions, while the formalism can be extended to interactions with neighboring locations as well.

The notion of success-driven motion is justified for the following reason: Comparing the term describing success-driven motion with a Fokker-Planck equation [43], one can conclude that it corresponds to a systematic drift with speed

\[
V(i,x,t) = \frac{\partial E(i,x,t)}{\partial x} \tag{26}
\]

i.e. individuals move into the direction of the gradient of the expected payoff, i.e. the direction of the (greatest)
increase of $E(i,x,t)$. The last term in Eq. (23) is a diffusion term which reflects effects of random motion in space [43]. It can be easily seen that, for $D_i > 0$, the diffusion term has a smoothing effect: It eventually reduces the proportion $p(i,x,t)$ in places $x$ where the second spatial derivative $\partial^2 p/\partial x^2$ is negative, in particular in places $x$ where the distribution $p(i,x,t)$ has maxima in space. In contrast, the proportion $p(i,x,t)$ increases in time, where $\partial^2 p/\partial x^2 > 0$, e.g. where the distribution has its minima.

### 3.1 Definition of Social Forces for Migration Games

When writing Eq. (26) explicitly, it becomes

$$V(i,x,t) = \sum_{j=1}^2 P_{ij} \frac{\partial p(j,x,t)}{\partial x} = \sum_{j=1}^2 f_{ij}(x,t).$$

(27)

Here, the expression

$$f_{ij}(x,t) = P_{ij} \frac{\partial p(j,x,t)}{\partial x}$$

(28)

(which can be extended by a saturation effects), may be interpreted as interaction force (“social force”) exerted by individuals using strategy $j$ on an individual using strategy $i$. It is visible that the sign of $P_{ij}$ determines the character of the force. The force is attractive for positive payoffs $P_{ij} > 0$ and repulsive for negative payoffs $P_{ij} < 0$. The direction of the force, however, is determined by spatial changes $\partial p(j,x,t)/\partial x$ in the strategy distribution $p(j,x,t)$ (i.e. not by the strategy distribution itself).

It is not the size of the payoffs $P_{ij}$ which determines the strength of the interaction force, but the payoff times the gradient of the distribution of the strategy $j$ one interacts with (and the availability and reachability of more favorable neighboring locations, if the saturation prefactor $C$ is taken into account). Due to the dependence on the gradient $\partial p(j,x,t)/\partial x$, the impact of a dispersed strategy $j$ on individuals using strategy $i$ is negligible. This particularly applies to scenarios with negative self-interactions ($P_{ij} < 0$).

Note that success-driven motion may be caused by repulsion away from the current location or by attraction towards more favorable neighborhoods. An interesting case is the game with the payoffs $P_{11} = P_{22} = -P$ and $P_{12} = P_{21} = Q > P$, where we have negative self-interactions among identical strategies and positive interactions between different strategies. Simulations for the no-imitation case show that, despite of the dispersive tendency of each strategy, strategies tend to agglomerate in certain locations thanks to the stronger attractive interactions between different strategies (see Fig. 3 in Ref. [42]).

The idea of social forces is long-standing. Montroll used the term to explain logistic growth laws [44], and Lewin introduced the concept of social fields to the social sciences in analogy to electrical fields in physics [45]. However, a formalization of a widely applicable social force concept was missing for a long time. In the meantime, social forces were successfully used to describe the dynamics of interacting vehicles [29] or pedestrians [27], but here the attractive or repulsive nature was just assumed. Attempts to systematically derive social forces from an underlying decision mechanism were based on direct pair interactions in behavioral spaces (e.g. opinion spaces), with the observation that imitative interactions or the readiness for compromises had attractive effects [19,46]. Here, for the first time, we present a formulation of social forces in game-theoretical terms. Considering the great variety of different games, depending on the respective specification of the payoffs $P_{ij}$, this is expected to find a wide range of applications, in particular as success-driven motion has been found to produce interesting and relevant pattern formation phenomena [42,39].

### 3.2 Linear Stability Analysis

In order to understand spatio-temporal pattern formation, it is not enough to formulate the (social) interaction forces determining the motion of individuals. We also need to grasp, why spatial patterns can emerge from small perturbations, even if the initial distribution of strategies is uniform (homogeneous) in space.

Assuming periodic boundary conditions (i.e. a circular space), we have $p(i,L,t) = p(i,0,t)$ and $\partial^2 p(i,L,t)/\partial x^2 = \partial^2 p(i,0,t)/\partial x^2$. By means of partial integration, one can prove that

$$\frac{\partial}{\partial t} \int_0^L dx \, p(i,x,t) = 0,$$

(29)

so that the normalization condition (19) is fulfilled. Furthermore, it can be shown that $p(i,x,t) \geq 0$ for all times $t$, if this is true for $t = 0$ for all strategies $i$ and locations $x$.

If Eq. (23) is written explicitly for $i = 1$, we get

$$\frac{\partial p(1,x,t)}{\partial t} = \left\{ P_{11} p(1,x,t) + P_{12} p(2,x,t) - P_{11} [p(1,x,t)]^2 - (P_{12} + P_{21}) p(1,x,t) p(2,x,t) - P_{22} [p(2,x,t)]^2 \right\} F_{12}(x,t) - \frac{\partial}{\partial x} \left\{ \left[ P_{11} \partial p(1,x,t)/\partial x \right. \right.$$

$$\left. + P_{12} \partial p(2,x,t)/\partial x \right\} F_{12}(x,t) + D_i \frac{\partial^2 p(1,x,t)}{\partial x^2},$$

(30)
and the equation for $i = 2$ looks identical, if only $p(1, x, t)$ and $p(2, x, t)$ are exchanged in all places, and the same is done with the indices 1 and 2. We will now assume a homogeneous initial condition $p(i, x, 0) = p(i, 0)$ (i.e. a uniform distribution of strategies $i$ in space) and study the spatio-temporal evolution of the deviations $\delta p(i, x, t) = p(i, x, t) - p(i, 0)$. Let us insert for $p(1, 0)$ one of the values $p_k(1)$, which are stationary solutions of the partial differential equation (23) due to $p(0, x, 0) = [1 - p(1, 0)]$ for $L = 1$. Assuming small deviations $\delta p(i, x, t)$ and linearizing Eq. (30) by neglecting non-linear terms, we obtain

$$\frac{\partial \delta p(1, x, t)}{\partial t} = \left\{ P_{11}p(1, 0) + P_{12}p(2, 0) - P_{11}[p(1, 0)]^2 \right\}$$

$$- (P_{12} + P_{21})p(1, 0)p(2, 0)$$

$$- 2P_{12}p(0, 0)\delta p(1, x, t) - 2P_{21}p(0, 0)\delta p(2, x, t)$$

$$- 2P_{12}p(2, 0)\delta p(1, x, t) + 2P_{21}p(2, 0)\delta p(2, x, t)$$

$$- p(1, 0)\left( P_{11} \frac{\partial^2 \delta p(1, x, t)}{\partial x^2} \right)$$

$$+ P_{12} \frac{\partial^2 \delta p(2, x, t)}{\partial x^2}$$

$$+ D_1 \frac{\partial^2 \delta p(1, x, t)}{\partial x^2}.$$  (31)

Again, a mutation term $W[\delta p(2, x, t) - \delta p(1, x, t)]$ reflecting spontaneous strategy changes may be added, see Eq. (13). The analogous equation for $\delta p(2, x, t)$ is obtained by exchanging strategies 1 and 2.

In Eq. (31), it can be easily seen that success-driven motion with $P_{ij} > 0$ has a similar functional form, but the opposite sign as the diffusion term. While the latter causes a homogenization in space, success-driven motion can cause local agglomeration [42], first of all for $P_{ij} > 0$.

It is known that linear partial differential equations like Eq. (31) are solved by (a superposition of) functions of the kind

$$\delta p(i, x, t) = e^{\lambda t} \left[ a_i \cos(kx) + b_i \sin(kx) \right],$$  (32)

where $a_i$ and $b_i$ are initial amplitudes, $\lambda = \lambda(k)$ is their growth rate (if positive, or a decay rate, if negative), and $k = k_n = 2\pi n/L$ with $n \in \{1, 2, \ldots\}$ are possible “wave numbers”. The “wave length” $2\pi/k = L/n$ may be imagined as the extension of a cluster of strategy $i$ in space. Obviously, possible wave lengths in case of a circular space of diameter $L$ are fractions $L/n$. The general solution of Eq. (31) is

$$\delta p(i, x, t) = \sum_{n=1}^{\infty} e^{\lambda(k_n)t} \left[ a_{i,n} \cos(k_n x) + b_{i,n} \sin(k_n x) \right],$$  (33)

i.e. a linear superposition of solutions of the form (32) with all possible wave numbers $k_n$. For $t = 0$, the exponential prefactor $e^{\lambda(k_n)t}$ becomes 1, and Eq. (33) may then be viewed as the Fourier series of the spatial dependence of the initial condition $\delta p(i, x, 0)$. Hence, the amplitudes $a_{i,n}$ and $b_{i,n}$ correspond to the Euler-Fourier coefficients [47].

Let us now determine the possible eigenvalues $\lambda(k)$. For the ansatz (32), we have $\partial \delta p(i, x, t)/\partial t = \lambda \delta p(i, x, t)$ and $\partial^2 \delta p(i, x, t)/\partial x^2 = -\kappa^2 \delta p(i, x, t)$. Therefore, the linearized equations can be cast into the following form of an eigenvalue problem with eigenvalues $\lambda$:

$$\begin{pmatrix} \delta p(1, x, t) \\ \delta p(2, x, t) \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \delta p(1, x, t) \\ \delta p(2, x, t) \end{pmatrix}.$$  (34)

Here, we have introduced the abbreviations

$$M_{11} = A_{11} + [p(1, 0)P_{11} - D_1]\kappa^2,$$

$$M_{12} = A_{12} + p(1, 0)P_{12}\kappa,$$

$$M_{21} = A_{21} + p(2, 0)P_{21}\kappa,$$

$$M_{22} = A_{22} + [p(2, 0)P_{22} - D_2]\kappa^2$$

with

$$A_{11} = P_{11}p(1, 0) + P_{12}p(2, 0) - P_{11}[p(1, 0)]^2$$

$$- (P_{12} + P_{21})p(1, 0)p(2, 0)$$

$$- 2P_{12}p(0, 0)\delta p(1, x, t) - 2P_{21}p(0, 0)\delta p(2, x, t)$$

$$- 2P_{12}p(2, 0)\delta p(1, x, t) + 2P_{21}p(2, 0)\delta p(2, x, t)$$

$$- p(1, 0)\left( P_{11} \frac{\partial^2 \delta p(1, x, t)}{\partial x^2} \right)$$

$$+ P_{12} \frac{\partial^2 \delta p(2, x, t)}{\partial x^2}$$

$$+ D_1 \frac{\partial^2 \delta p(1, x, t)}{\partial x^2}.$$  (31)

The eigenvalue problem (34) can only be solved, if the determinant of the matrix $(M - \lambda I)$ vanishes, where $I$ denotes the unit matrix [47]. In other words, $\lambda$ are solutions of the so-called “characteristic polynomial”

$$\left( M_{11} - \lambda \right)\left( M_{22} - \lambda \right) - M_{12}M_{21} = 0.$$  (40)

This polynomial is of degree 2 in $\lambda$ and has the following two solutions:

$$\lambda = \frac{M_{11} + M_{22}}{2} \pm \frac{1}{2} \sqrt{(M_{11} + M_{22})^2 - 4(M_{11}M_{22} - M_{12}M_{21})}.$$  (41)

Now, we will focus on the particularly interesting case, where the mathematical expression under the root is nonnegative (but the case of a negative value if $4M_{12}M_{21} <$
\[-(M_{11} - M_{22})^2\] could be treated as well). It can be shown that \(\lambda\) becomes positive, if one of the following instability conditions is fulfilled:

\[
M_{11} + M_{22} > 0 \tag{42}
\]
or

\[
M_{11}M_{22} < M_{12}M_{21} \tag{43}
\]

In this case, we expect the amplitudes of the small deviations \(\delta p(i, x, t)\) to grow over time, which gives rise to spatial pattern formation (such as segregation). Inserting the abbreviations (35) to (38), the instability conditions become

\[
\left\{ A_{11} + [p(1,0)P_{11} - D_1]\kappa^2 \right\} \\
+ \left\{ A_{22} + [p(2,0)P_{22} - D_2]\kappa^2 \right\} > 0 \tag{44}
\]

and

\[
\left\{ A_{11} + [p(1,0)P_{11} - D_1]\kappa^2 \right\} \\
\times \left\{ A_{22} + [p(2,0)P_{22} - D_2]\kappa^2 \right\} \\
< \left[ A_{12} + p(1,0)P_{12}\kappa^2 \right] \left[ A_{21} + p(2,0)P_{21}\kappa^2 \right]. \tag{45}
\]

If \(\kappa\) is large enough (i.e. if the related cluster size is sufficiently small), the terms resulting from the right-hand side of the replicator equation (4) become negligible, and the inequalities (45), (44) simplify to

\[
p(1,0)P_{11} + p(2,0)P_{22} > D_1 + D_2 \tag{46}
\]

and

\[
[p(1,0)P_{11} - D_1][p(2,0)P_{22} - D_2] < p(1,0)P_{12}p(2,0)P_{21}. \tag{47}
\]

Roughly speaking (neglecting diffusion effects), Eq. (47) says that the interactions between different strategies (measured by the product of the related payoffs \(P_{ij}\) and \(P_{jk}\)) must be stronger than the interactions among identical strategies (measured by the product of the related payoffs \(P_{ij}\) and \(P_{ki}\)). Before, this instability condition has been studied in the context of success-driven motion without imitation and for games with symmetrical payoff matrices (i.e. \(P_{ij} = P_{ji}\)), which show a particular behavior [48]. Over here, in contrast, we investigate continuous spatial games involving imitation (selection of more successful strategies), and focus on asymmetrical games such as the prisoner’s dilemma and the snowdrift game (see Sec. 4), which behave very differently.

Generally speaking, if condition (47) is fulfilled, we expect emergent spatio-temporal pattern formation, specifically the self-organization of cooperative clusters, and a related increase in the level of cooperation. When spatial interactions are neglected, we stay with Eq. (11) and cannot find this favorable pattern-formation effect.

### 4 Results for the Prisoner’s Dilemma and the Snowdrift Game

Let us now discuss a variety of different cases:

1. In case of diffusive motion only (i.e. no success-driven motion), Eq. (46) must be replaced by

\[
0 > D_1 + D_2 \tag{48}
\]

and (47) by

\[
D_1D_2 < 0, \tag{49}
\]

which cannot be fulfilled. Therefore, diffusion without success-driven motion cannot support cooperation.

2. In the case of success-driven motion with \(P_{11} = R\), \(P_{12} = S\), \(P_{21} = T\), and \(P_{22} = P\), Eq. (47) becomes

\[
[p(1,0)R - D_1][p(2,0)P - D_2] < p(1,0)p(2,0)ST, \tag{50}
\]

while Eq. (46) implies

\[
p(1,0)R + p(2,0)P > D_1 + D_2. \tag{51}
\]

Considering \(p(2,0) = 1 - p(1,0)\), this gives

\[
[p(1,0)R - D_1][\{1 - p(1,0)\}P - D_2] < p(1,0)[1 - p(1,0)]ST, \tag{52}
\]

and

\[
p(1,0)R + [1 - p(1,0)]P > D_1 + D_2. \tag{53}
\]

These instability conditions hold for both, the prisoner’s dilemma and the snowdrift game. Note, however, that condition (52) is not invariant with respect to shifts of all payoffs \(P_{ij}\) by a constant value \(c\), in contrast to the replicator equations (9).

3. In case of the prisoner’s dilemma, the stable stationary solution for the case without spontaneous strategy changes is \(p(1,0) = p_1(1) = 0\), which simplifies the instability conditions further, yielding

\[
-D_1(P - D_2) < 0, \tag{54}
\]

and

\[
P > D_1 + D_2. \tag{55}
\]

In order to fulfill one of these conditions, the punishment \(P\) must be large enough to support pattern formation. Specifically, we require \(P > D_2 \geq 0\).

4. In the special case \(P = S = 0\) discussed in Ref. [39], a finite value of \(p(1,0)\) is needed (which can be easily reached by spontaneous strategy changes): One can show that Eqs. (52) and (53) become

\[
-|p(1,0)R - D_1|D_2 < 0 \tag{56}
\]

and

\[
p(1,0)R > D_1 + D_2, \tag{57}
\]

which requires \(p(1,0)R > D_1\).
5. Neglecting diffusion for a moment (i.e. setting $D_1 = D_2 = 0$), no pattern formation should occur, if the condition

$$RP > ST$$

(58)

and, at the same time,

$$p(1,0)R + [1 - p(1,0)]P < 0$$

(59)

is fulfilled. Equation (58) implies the stability condition

$$S < \frac{RP}{T}.$$ 

(60)

Besides $T > R$, we have to consider here that $S < P$ in the prisoner’s dilemma and $S > P$ in the snowdrift game (see Fig. 1). Of course, we also need to take into account Eq. (59), which implies $P < 0$ for the stationary solution $p(1,0) = p_1(1) = 0$ of the prisoner’s dilemma and generally

$$P < -\frac{p(1,0)R}{1 - p(1,0)}.$$ 

(61)

In case of the snowdrift game [49], it is adequate to insert the stationary solution $p(1,0) = p_3(1) = A/B$, which is stable in case of no spatial interactions. This leads to the condition

$$P < -\frac{p_3(1)R}{1 - p_3(1)} = \frac{AR}{A - B} = \frac{(P - S)R}{T - R},$$

(62)

that is

$$SR < (2R - T)P.$$ 

(63)

The question is, whether this condition will reduce the previously determined area of stability given by $S < RP/T$ with $P < 0$, see Eq. (60) and Fig. 1. This would be the case, if $(2 - T/R)P > RP/T$, which by multiplication with $RT/P$ becomes $(2RT - T^2) > R^2$ or $(T - R)^2 < 0$. Since this condition cannot be fulfilled, it does not impose any further restrictions on the stability area in the snowdrift game.

6. Finally, let us assume that the stability conditions $ST - RP < 0$ and $p(1,0)R + [1 - p(1,0)]P < 0$ for the case without diffusion ($D_1 = D_2 = 0$) are fulfilled, so that no patterns will emerge. Then, depending on the parameter values, the instability condition following from Eq. (52),

$$-(D_1 - D_2)\frac{[1 - p(1,0)]}{0} P$$

$$< D_2 \frac{[p(1,0)R + [1 - p(1,0)]P]}{0 < 0}$$

$$-D_1\frac{D_2 + p(1,0)[1 - p(1,0)]}{\geq 0} (ST - RP),$$

(64)

may still be matched, if $(D_1 - D_2)$ is sufficiently large. Therefore, asymmetrical diffusion $(D_1 \neq D_2)$ can trigger a pattern formation instability, where the spatio-temporal strategy distribution without diffusion would be stable. The situation is clearly different for symmetrical diffusion with $D_1 = D_2$, which cannot support pattern formation.

![Fig. 1. Payoff-dependence of pattern-formation in the prisoner’s dilemma with $S < P$ and the snowdrift game with $S > P$ according to a linear stability analysis for spatial games with success-driven motion and no diffusion. One can clearly see that pattern-formation prevails (green area), and that there is only a small area for $P < 0$ (marked red), where a homogeneous initial condition is stable with respect to perturbations or fluctuations.](image)

5 Summary and Outlook

In this paper, we have started from the game-dynamical equations (replicator equation), which can be derived from imitative pair interactions between individuals [17,18]. It has been shown that zero cooperation is expected in the prisoner’s dilemma, if no spontaneous strategy mutations are taken into account, otherwise there will be a finite, but usually low level of cooperation. In the snowdrift game, in contrast, the stationary solution corresponding to no cooperation is unstable, and there is a stable solution with a finite level of cooperation.

These introductory considerations have been primarily performed to illustrate the major differences caused by the introduction of spatial interactions based on success-driven motion and diffusion. While diffusion itself tends to support homogeneous strategy distributions rather than pattern formation, success-driven motion implies an unstable spatio-temporal dynamics under a wide range of conditions. As a consequence, small fluctuations can destabilize a homogeneous distribution of strategies. Under such conditions, the formation of emergent patterns is expected. The resulting dynamics may be understood in terms of social forces, which have been formulated here in game-theoretical terms.

For the case without imitation of superior strategies and symmetrical payoffs ($P_{ij} = P_{ji}$), it has been shown [48] that the instability conditions predict the results of
agent-based simulations with a discretized version of the above proposed model surprisingly well. This is also expected to be true for the non-symmetrical games studied here, in particular as we found that the influence of imitation on the instability condition is negligible, if the wave number \( k \) characterizing inhomogeneities in the initial distribution is large. This simplified the stability analysis a lot. Moreover, it was shown that asymmetrical diffusion can drive our game-theoretical model with success-driven motion from the stable regime into the unstable regime, which reminds of the Turing instability [50–52] and noise-induced transitions [53]. However, the underlying equations are very different, as is elaborated in Ref. [48].

Finally, note that migration may be considered as one realization of success-driven motion. Before, the statistics of migration behavior was modeled by the gravity law [54,55] or entropy approaches [56,57], while the dynamics was described by partial differential equations [58,59] and models from statistical physics [60]. The particular potential of the approach proposed in this paper lies in the integration of migration into a game-theoretical framework, as we formalize success-driven motion in terms of payoffs and strategy distributions in space and time. Such integrated approaches are needed in the social sciences to allow for consistent interpretations of empirical findings within a consistent framework.

Acknowledgments

The author would like to thank Peter Felten for preparing Fig. 1 and Christoph Hauert for comments on manuscript [39].

References

1. J. von Neumann, O. Morgenstern, Theory of Games and Economic Behavior (Princeton University Press, Princeton, 1944).
2. R. Axelrod, The Evolution of Cooperation (Basic, New York, 1984).
3. A. Rapoport, Game Theory as a Theory of Conflict Resolution (Reidel, Dordrecht, 1974).
4. H. Gintis, Game Theory Evolving (Princeton University Press, Princeton, 2000).
5. D. Fudenberg, J. Tirole, Game Theory (MIT Press, Cambridge, MA, 1991).
6. K. Binmore, Playing for Real (Oxford University Press, Oxford, 2007).
7. M. A. Nowak, Evolutionary Dynamics (Belknap Press, Cambridge, MA, 2006).
8. M. J. Osborne, An Introduction to Game Theory (Oxford University Press, New York, 2004).
9. S.-C. Kolm, J. M. Ythier, Handbook of the Economics of Giving, Altruism, and Reciprocity, Vol. I-II (North-Holland, Amsterdam, 2006).
10. G. Szabó, G. Fath, Evolutionary games on graphs. Phys. Rep. 446, 97-216 (2007).
11. J. H. Kagel, A. E. Roth (eds.), Handbook of Experimental Economics (Princeton University Press, Princeton, 1995).
12. C. F. Camerer, Behavioral Game Theory (Princeton University Press, Princeton, 2003).
13. J. Heinrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis (eds.) Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies (Oxford University Press, Oxford, 2004).
14. U. Fischbacher, S. Gächter, E. Fehr, Are people conditionally cooperative? Evidence from a public goods experiment. Economics Letters 71, 397–404 (2001).
15. E. Sober, D. S. Wilson, Unto Others: The Evolution and Psychology of Unselfish Behavior (Harvard Univ. Press, Cambridge, Massachusetts, 1998).
16. M. W. Macy, A. Flache, Learning dynamics in social dilemmas. Proc. Natl. Acad. Sci. (USA) 99, Suppl. 3, 7229-7236 (2002).
17. D. Helbing, A mathematical model for behavioral changes by pair interactions, in Economic Evolution and Demographic Change. Formal Models in Social Sciences, edited by G. Haag, U. Mueller, K. G. Troitzsch (Springer, Berlin, 1992), pp. 330-348.
18. D. Helbing, A stochastic behavioral model and a ‘microscopic’ foundation of evolutionary game theory. Theory and Decision 40, 149-179 (1996).
19. D. Helbing, Quantitative Sociodynamics (Kluwer Academic, Dordrecht, 1995).
20. J. Hofbauer, K. Sigmund, Evolutionstheorie und dynamische Systeme (Paul Parey, Berlin, 1984).
21. J. Hofbauer and K. Sigmund, The Theory of Evolution and Dynamical Systems (Cambridge University, Cambridge, 1988).
22. J. W. Weibull, Evolutionary Game Theory (MIT Press, Cambridge, MA, 1995).
23. L. Samuelson, Evolutionary Games and Equilibrium Selection (MIT Press, Cambridge, MA, 1997).
24. M. Eigen, The selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften 58, 465f (1971).
25. R. A. Fisher, The Genetical Theory of Natural Selection (Oxford University Press, Oxford, 1930).
26. M. Eigen, P. Schuster, The Hypercycle (Springer, Berlin, 1979).
27. D. Helbing, P. Molnár, Social force model for pedestrian dynamics. Phys. Rev. E 51, 4282-4286 (1995).
28. D. Helbing, I. Farkas, T. Vicsek, Simulating dynamical features of escape panic. Nature 407, 487-490 (2000).
29. D. Helbing and B. Tilch (1998) Generalized force model of traffic dynamics. Physical Review E 58, 133-138.
30. D. Helbing, A. Johansson, H. Z. Al-Abideen, The dynamics of crowd disasters: An empirical study. Phys. Rev. E 75, 046109 (2007).
31. D. Helbing, L. Buzna, A. Johansson, T. Werner, Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. Transportation Science 39(1), 1-24 (2005).
32. S. P. Hoogendoorn, W. Daamen, Pedestrian behavior at bottlenecks. Transportation Science 39(2), 147-159 (2005).
33. A. Johansson, D. Helbing, P.S. Shukla, Specification of the social force pedestrian model by evolutionary adjustment to video tracking data. Advances in Complex Systems 10, 271-288 (2007).
34. W. Yu, A. Johansson, Modeling crowd turbulence by many-particle simulations. Phys. Rev. E 76, 046105 (2007).
35. J. W. Pepper, Simple models of assortment through environmental feedback. *Artificial Life* 13, 1-9 (2007).
36. C. Hauert, S. De Monte, J. Hofbauer, K. Sigmund, Volunteering as red queen mechanism for cooperation in public goods games. Science 296, 1129-1132 (2002).
37. M. Doebeli, C. Hauert, Models of cooperation based on the prisoner’s dilemma and the snowdrift game, *Ecology Letters* 8, 748-766 (2005).
38. C. Hauert, Fundamental clusters in spatial 2x2 games. *Proc. R. Soc. Lond. B* 268, 761-769 (2001).
39. D. Helbing, W. Yu, Migration as a mechanism to promote cooperation. *Advances in Complex Systems*, in print (2008).
40. G. Szabó, C. Hauert, Phase transitions and volunteering in spatial public goods games. Phys. Rev. Lett. 89, 118101 (2002).
41. D. Helbing, T. Vicsek, Optimal self-organization. New Journal of Physics 1, 13.1-13.17 (1999).
42. D. Helbing, T. Platkowski, Self-organization in space and induced by fluctuations, International Journal of Chaos Theory and Applications 5(4), 47-62 (2000).
43. H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications* (Springer, Berlin, 1989).
44. E. W. Montroll, Social dynamics and the quantifying of social forces. Proc. Natl. Acad. Sci USA 75, 4633-4637 (1978).
45. K. Lewin, *Field Theory in the Social Science* (Harper & Brothers, New York, 1951).
46. D. Helbing, A mathematical model for the behavior of individuals in a social field. Journal of Mathematical Sociology 19 (3), 189-219 (1994).
47. K. F. Riley, M. P. Hobson & S. J. Bence, *Mathematical Methods for Physics and Engineering* (Cambridge University, Cambridge, 2006).
48. D. Helbing, T. Platkowski, Drift- or fluctuation-induced ordering and self-organisation in driven many-particle systems. Europhysics Letters 60, 227-233 (2000).
49. C. Hauert, M. Doebeli, Spatial structure often inhibits the evolution of cooperation in the snowdrift game. Nature 428, 643-646 (2004).
50. A. M. Turing, The chemical basis of morphogenesis. Philosophical Transactions of the Royal Society of London B 237, 37-72 (1952).
51. J. D. Murray, *Lectures on Nonlinear Differential Equation-Models in Biology* (Claderon Press, Oxford, 1977).
52. A. D. Kessler, H. Levine, Fluctuation-induced diffusive instabilities. Nature 394, 556-558 (1998).
53. W. Horsthemke, R. Lefever, *Noise-Induced Transitions* (Springer, Berlin, 1984).
54. E. Ravenstein, The birthplaces of the people and the laws of migration. The Geographical Magazine III, 173-177, 201-206, 229-233 (1876).
55. G. K. Zipf, The P1P2/D hypothesis on the intercity movement of persons. Sociological Review 11, 677-686 (1946).
56. A. G. Wilson, Entropy maximising models in the theory of trip distribution, mode split and route split. J. Transport Econ. Policy 3, 108-126 (1969).
57. S. Brice, Derivation of nested transport models within a mathematical programming framework. Transpn. Res. B 23(1), 19-28 (1989).
58. H. Hotelling, A mathematical theory of migration. Environment and Planning A 10, 1223-1239 (1978).
59. T. Puu, A simplified model of spatiotemporal population dynamics. Environment and Planning A 17, 1263-1269 (1985).
60. W. Weidlich, G. Haag (eds.) *Interregional Migration* (Springer, Berlin, 1988).